Access DB# 35506

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner # : _____ Date: _____
Art Unit: _____ Phone Number 30_____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

***For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.**

FOR OFFICIAL USE ONLY

****

STAFF USE ONLY

Searcher: Bostick
Searcher Phone #: 308-4477
Searcher Location: CM1-1E17
Date Searcher Picked Up: 2/15/01
Date Completed: 3/1/01
Searcher Prep & Review Time: _____
Clerical Prep Time: 3
Online Time: _____

Type of Search

NA Sequence (#) 2
AA Sequence (#) 2
Structure (#) _____
Bibliographic _____
Litigation _____
Fulltext _____
Patent Family _____
Other _____

Vendors and cost where applicable

STN _____
Dialog _____
Questel/Orbit _____
Dr.Link _____
Lexis/Nexis _____
Sequence Systems abSS02
WWW/Internet _____
Other (specify) _____

PTO-1590 (1-2000)

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 19:53:28 ; Search time 119.29 Seconds
(without alignments)
2220.154 Million cell updates/sec

Title: US-09-320-713-1
Perfect score: 705
Sequence: 1 ggcacgagtggacacggatg.....aaaaaaaaaaaaaaaaaaaa 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	704.2	99.9	705	21	Z36834	Partial nucleotide
2	698.6	99.1	1067	21	Z36836	Nucleotide sequenc
3	640.4	90.8	1047	21	Z29728	Human Interleukin
4	256.4	36.4	591	21	A09153	Human interleukin-
5	230	32.6	230	21	Z29731	Virtual DNA fragme
6	141.2	20.0	332	21	Z36856	cDNA clone HTGED19
7	67	9.5	1453	14	Q50573	Asparaginylendopep
8	67	9.5	1640	14	Q50575	Asparaginylendopep
9	67	9.5	1910	14	Q50579	Asparaginylendopep
10	65.8	9.3	2010	21	A23428	cDNA encoding huma
11	61.4	8.7	1554	20	X16675	Xenopus WA545 prot
12	61.2	8.7	882	18	T72173	Alzheimer's diseas
13	61.2	8.7	4237	19	V61487	Human secreted pro
14	61	8.7	1181	19	V59803	Human secreted pro
15	61	8.7	1212	19	V59686	Human secreted pro
16	60.6	8.6	1135	20	X80676	Clone yb7_1 encodi
17	59.2	8.4	281	21	A00286	Human colon cancer
18	59	8.4	1527	21	A16633	Human secreted pro
19	58.8	8.3	1037	21	A26447	Human secreted pro
20	58.6	8.3	1740	21	Z36914	DNA encoding an ac
21	58.6	8.3	1801	21	Z36893	cDNA encoding an a
22	58.4	8.3	2665	19	V32592	Schwannomin-bindin
23	58.2	8.3	491	21	Z52570	Human secreted pro
24	58.2	8.3	1602	18	T43204	Human mitogen-acti
25	58.2	8.3	1602	20	X07066	Human mitogen acti
26	58	8.2	550	19	V41916	Nucleotide sequenc
27	58	8.2	1212	20	Z28298	Rat neuronal immed
28	58	8.2	1231	20	Z28265	Rat neuronal immed
29	57.8	8.2	1963	21	Z98097	Human secreted pro
30	57.6	8.2	997	20	X52274	Protein PRO244 cDN
31	57.6	8.2	1558	17	T28255	Survival motor neu
32	57.6	8.2	1560	17	T18828	Human survival mot
33	57.6	8.2	1582	17	T28259	Survival motor neu
34	57.6	8.2	1582	17	T18831	Human survival mot
35	57.6	8.2	2214	19	V40524	Homo sapiens CO122
c 36	57.4	8.1	467	20	V64423	Mouse developing l
37	57.4	8.1	916	18	T65651	cDNA encoding anti
38	57.4	8.1	916	19	V62000	R. prolixus NO-r p
39	57.2	8.1	1410	12	Q15022	Hyoscyamine 6 beta
40	57.2	8.1	6200	21	Z49493	Human alpha-L-idur
41	57	8.1	615	18	V02881	Human HMGI-C aberr
42	57	8.1	1701	21	Z64982	Membrane-bound pro
43	56.8	8.1	1965	21	Z98033	Human secreted pro
44	56.8	8.1	2136	21	Z65065	Membrane-bound pro
45	56.6	8.0	2300	20	Z00686	Human GPC3 DNA. H

ALIGNMENTS

RESULT 1
Z36834
ID Z36834 standard; cDNA; 705 BP.
XX
AC Z36834;
XX
DT 13-MAR-2000 (first entry)
XX
DE Partial nucleotide sequence of human interleukin-21.
XX
KW Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW immune cell chemotaxis; haematopoietic cell disorder;
KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW inflammation; hyperproliferative disorder; tissue regeneration;
KW embryonic stem cell differentiation; embryonic stem cell proliferation;
KW haematopoietic lineage; allergic asthma; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 2..265
FT /*tag= a
FT /product= "partial IL-21"
XX
PN WO9961617-A1.
XX
PD 02-DEC-1999.
XX
PF 27-MAY-1999; 99WO-US11644.
XX
PR 29-MAY-1998; 98US-0087340.
PR 10-SEP-1998; 98US-0099805.
PR 30-APR-1999; 99US-0131965.
XX

```
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
DR   P-PSDB; Y53890.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Claim 10; Fig 1; 170pp; English.
XX
CC   The present sequence encodes a partial human interleukin-21 (IL-21)
CC   protein. The specification also describes IL-22 polynucleotides and
CC   polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC   of apoptotic T-cells. IL-21 and IL-22 may be useful in treating
CC   deficiencies or disorders of the immune system, by activating or
CC   inhibiting the proliferation, differentiation, or mobilization
CC   (chemotaxis) of immune cells, treating or detecting deficiencies
CC   of haematopoietic cells, to modulate haemostatic or thrombolytic
CC   activity, in treating or detecting autoimmune disorders, treating
CC   asthma (particularly allergic asthma) or other respiratory problems,
CC   to treat and/or prevent organ rejection or graft-versus-host disease
CC   (GVHD), to modulate inflammation, to treat or detect hyperproliferative
CC   disorders, to treat or detect infectious agents, to differentiate,
CC   proliferate and attract cells, leading to the regeneration of tissues,
CC   IL-21 and IL-22 may also increase or decrease the differentiation or
CC   proliferation of embryonic stem cells and haematopoietic lineage, may
CC   be used to modulate mammalian characteristics.
XX
SQ   Sequence 705 BP; 156 A; 226 C; 169 G; 152 T; 2 other;


  Query Match             99.9%;  Score 704.2;  DB 21;  Length 705;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-126;
  Matches  705;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ggcacgagtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcct 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ggcacgagtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcct 60

Qy     61 gtgcagaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gtgcagaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcg 120

Qy    121 gctgctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggct 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gctgctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggct 180

Qy    181 ccccacacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcac 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ccccacacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcac 240

Qy    241 ctgcgtgctgccccgttcagtgtgaccgccaaggccgtggggcccttagactggacacgt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ctgcgtgctgccccgttcagtgtgaccgccaaggccgtggggcccttagactggacacgt 300

Qy    301 gtgctccccagagggcacccccctatttatgtgtatttattgttatttatatgcctccccc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gtgctccccagagggcacccccctatttatgtgtatttattgttatttatatgcctccccc 360

Qy    361 aacactacccttggggtctgggcattccccgtgtctggaggacagcccccccactgttctc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 aacactacccttggggtctgggcattccccgtgtctggaggacagcccccccactgttctc 420

Qy    421 ctcatctccagcctcagtagttgggggtwgaaggagctcagcacctcttccagcccttaa 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ctcatctccagcctcagtagttgggggtwgaaggagctcagcacctcttccagcccttaa 480

Qy    481 agctgcagaaaaggtgtcacacggctgcctgtaccttggytccctgtcctgctcccggct 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 agctgcagaaaaggtgtcacacggctgcctgtaccttggytccctgtcctgctcccggct 540

Qy    541 tcccttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttgga 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 tcccttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttgga 600

Qy    601 agtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacac 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 agtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacac 660

Qy    661 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
          |||||||||||||||||||||||||||||||||||||||||||||
Db    661 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705


RESULT   2
Z36836
ID   Z36836 standard; cDNA; 1067 BP.
XX
AC   Z36836;
XX
DT   13-MAR-2000  (first entry)
XX
DE   Nucleotide sequence of human interleukin-21.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             34..627
FT                   /*tag=  a
FT                   /product= "IL-21"
FT   sig_peptide     34..1060
FT                   /*tag=  b
FT   sig_peptide     34..87
FT                   /*tag=  c
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
DR   P-PSDB; Y53892.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Claim 8; Fig 6A-B; 170pp; English.
XX
CC   The present sequence encodes a human interleukin-21 (IL-21)
CC   protein. The specification also describes IL-22 polynucleotides and
CC   polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC   of apoptotic T-cells. IL-21 and IL-22 may be useful in treating
CC   deficiencies or disorders of the immune system, by activating or
CC   inhibiting the proliferation, differentiation, or mobilization
CC   (chemotaxis) of immune cells, treating or detecting deficiencies
```

```
CC   of haematopoietic cells, to modulate haemostatic or thrombolytic
CC   activity, in treating or detecting autoimmune disorders, treating
CC   asthma (particularly allergic asthma) or other respiratory problems,
CC   to treat and/or prevent organ rejection or graft-versus-host disease
CC   (GVHD), to modulate inflammation, to treat or detect hyperproliferative
CC   disorders, to treat or detect infectious agents, to differentiate,
CC   proliferate and attract cells, leading to the regeneration of tissues,
CC   IL-21 and IL-22 may also increase or decrease the differentiation or
CC   proliferation of embryonic stem cells and haematopoietic lineage, may
CC   be used to modulate mammalian characteristics.
XX
SQ   Sequence 1067 BP; 215 A; 371 C; 273 G; 206 T; 2 other;


  Query Match             99.1%;  Score 698.6;  DB 21;  Length 1067;
  Best Local Similarity   99.9%;  Pred. No. 1e-124;
  Matches  700;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 cgagtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgc 64
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     367 cgggtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgc 426

Qy      65 agaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctg 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     427 agaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctg 486

Qy     125 ctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctcccc 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     487 ctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctcccc 546

Qy     185 acacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacctgc 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     547 acacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacctgc 606

Qy     245 gtgctgccccgttcagtgtgaccgccaaggccgtggggcccttagactggacacgtgtgc 304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     607 gtgctgccccgttcagtgtgaccgccaaggccgtggggcccttagactggacacgtgtgc 666

Qy     305 tccccagagggcacccccctatttatgtgtatttattgttatttatatgcctcccccaaca 364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     667 tccccagagggcacccccctatttatgtgtatttattgttatttatatgcctcccccaaca 726

Qy     365 ctacccttggggtctgggcattccccgtgtctggaggacagccccccactgttctcctca 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     727 ctacccttggggtctgggcattccccgtgtctggaggacagccccccactgttctcctca 786

Qy     425 tctccagcctcagtagttgggggtwgaaggagctcagcacctcttccagcccttaaagct 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     787 tctccagcctcagtagttgggggtwgaaggagctcagcacctcttccagcccttaaagct 846

Qy     485 gcagaaaaggtgtcacacggctgcctgtaccttggytccctgtcctgctcccggcttccc 544
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     847 gcagaaaaggtgtcacacggctgcctgtaccttggytccctgtcctgctcccggcttccc 906

Qy     545 ttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttggaagta 604
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     907 ttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttggaagta 966

Qy     605 cccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaa 664
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     967 cccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaa 1026

Qy     665 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||||||||
Db    1027 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1067


RESULT   3
Z29728
ID   Z29728 standard; cDNA; 1047 BP.
XX
AC   Z29728;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Human Interleukin 17C, PRO1122 cDNA.
XX
KW   Interleukin; IL-17C; PRO1122 polypeptide; clone DNA62377-1381-1; UNQ561;
KW   cytokine IL-17; cytotoxic T-lymphocyte-associated antigen 8; CTLA-8;
KW   hybridisation probe; antagonist; degenerative cartilaginous disorder;
KW   agonist; diagnose; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     50..103
FT                   /*tag=  a
FT   mat_peptide     50..634
FT                   /*tag=  b
FT                   /label= Mature_Interleukin_17C_polypeptide
FT                   /note= "Designated as clone DNA62377-1381-1"
XX
PN   WO9960127-A2.
XX
PD   25-NOV-1999.
XX
PF   14-MAY-1999;   99WO-US10733.
XX
PR   15-MAY-1998;   98US-0085579.
PR   23-DEC-1998;   98US-0113621.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Filvaroff E,  Goddard A,  Gurney AL,  Li H,  Wood WI;
XX
DR   WPI; 2000-116314/10.
DR   P-PSDB; Y44460.
XX
PT   New polypeptides designated PRO1031 and PRO1122 used to treat a
PT   degenerative cartilaginous disorder -
XX
PS   Claim 2; Fig 4; 141pp; English.
XX
CC   The present sequence is the cDNA clone DNA62377-1381-1, encoding the
CC   human PRO1122 polypeptide, also referred to as UNQ561, and as
CC   interleukin-17C (IL-17C). This sequence has identity with the
CC   cytokine IL-17 and cytotoxic T-lymphocyte-associated antigen 8 (CTLA-8).
CC   PRO1122 is expressed in pancreas, small intestine, stomach and testis
CC   also. It shares about 26-28% amino acid identity with IL-17 and IL-17B.
CC   The entire coding region of IL-17C can be used as hybridisation probe.
CC   The PRO1122 polypeptide, agonist or antagonist, is used to diagnose and
CC   treat a degenerative cartilaginous disorder.
XX
SQ   Sequence 1047 BP; 177 A; 381 C; 280 G; 209 T; 0 other;


  Query Match             90.8%;  Score 640.4;  DB 21;  Length 1047;
  Best Local Similarity   98.3%;  Pred. No. 1.2e-113;
  Matches  655;  Conservative    2;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       5 cgagtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgc 64
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 cgtgtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgc 442

Qy      65 agaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctg 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 agaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctg 502

Qy     125 ctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctcccc 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 ctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctcccc 562

Qy     185 acacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacctgc 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 acacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacctgc 622
```

```
Qy    245 gtgctgccccgttcagtgtgaccgccaaggccgtggggcccttagactggacacgtgtgc 304
          ||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||
Db    623 gtgctgccccgttcagtgtgaccgccgaggccgtggggcccctagactggacacgtgtgc 682

Qy    305 tccccagagggcacccccctatttatgtgtatttattgttatttatatgcctcccccaaca 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 tccccagagggcacccccctatttatgtgtatttattgttatttatatgcctcccccaaca 742

Qy    365 ctacccttggggtctgggcattccccgtgtctggaggacagccccccactgttctcctca 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 ctacccttggggtctgggcattccccgtgtctggaggacagccccccactgttctcctca 802

Qy    425 tctccagcctcagtagttgggggtwgaaggagctcagcacctcttccagcccttaaagct 484
          ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db    803 tctccagcctcagtagttgggggtagaaggagctcagcacctcttccagcccttaaagct 862

Qy    485 gcagaaaaggtgtcacacggctgcctgtaccttggytccctgtcctgctcccggcttccc 544
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db    863 gcagaaaaggtgtcacacggctgcctgtaccttggctccctgtcctgctcccggcttccc 922

Qy    545 ttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttggaagta 604
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    923 ttaccctatcactggcctcagg-ccccgcaggctgcctcttcccaacctccttggaagta 981

Qy    605 cccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaa 664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 cccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacatcc 1041

Qy    665 aaaaaa 670
              ||||
Db   1042 ccaaaa 1047
```

RESULT 4
A09153
ID A09153 standard; DNA; 591 BP.
XX
AC A09153;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human interleukin-17 (IL-17) homologue coding sequence.
XX
KW Interleukin 17; IL-17; haematopoiesis; chemotherapy; cytostatic;
KW antianemic; cardiant; hemostatic; anti-inflammatory; anti-HIV; ss.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	sig_peptide	1..54
FT		/*tag= a
FT	mat_peptide	55..591
FT		/*tag= b

XX
PN WO200020593-A1.
XX
PD 13-APR-2000.
XX
PF 30-SEP-1999; 99WO-US22678.
XX
PR 02-OCT-1998; 98US-0102883.
PR 01-DEC-1998; 98US-0110405.
PR 11-JUN-1999; 99US-0138910.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Glasebrook AL, Su EW, Wei J, Liu L;
XX
DR WPI; 2000-303778/26.
DR P-PSDB; Y92238.
XX
PT Nucleic acid encoding an interleukin-17 (IL-17) homolog polypeptide
PT which enhances hematopoiesis, useful for treating e.g. anemia,
PT thrombocytopenia, viral and bacterial infections
XX
PS Claim 6; Page 92; 111pp; English.
XX
CC Interleukin 17 (IL-17) stimulates hematopoiesis and production of
CC neutrophils, granulocytes, or platelets, this may be useful during
CC chemotherapy. IL-17 homologues have at least one activity selected
CC from induction of cytotoxic T cells, induction of lymphokine-activated
CC killer cell proliferation or a B or T cell stimulation. The IL-17
CC homologue may also be used to treat viral or bacterial infections,
CC immune related diseases, anemia, leukemia, thrombocytopenia, uremia,
CC Von Willebrand disease, postoperative cardiovascular dysfunction,
CC treatment of AIDS (acquired immune deficiency syndrome)-related bone
CC marrow failure, and inflammatory diseases of the gastrointestinal
CC system, joints, and lungs.
XX
SQ Sequence 591 BP; 91 A; 223 C; 179 G; 98 T; 0 other;

Query Match 36.4%; Score 256.4; DB 21; Length 591;
Best Local Similarity 99.6%; Pred. No. 1.1e-40;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      5 cgagtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgc 64
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 cgtgtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgc 393

Qy     65 agaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctg 124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 agaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctg 453

Qy    125 ctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctcccc 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 ctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctcccc 513

Qy    185 acacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacctgc 244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 acacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacctgc 573

Qy    245 gtgctgccccgttcagtg 262
          ||||||||||||||||||
Db    574 gtgctgccccgttcagtg 591
```

RESULT 5
Z29731
ID Z29731 standard; DNA; 230 BP.
XX
AC Z29731;
XX
DT 27-MAR-2000 (first entry)
XX
DE Virtual DNA fragment DNA49665, for isolation of human PRO1122 cDNA.
XX
KW Interleukin; IL-17C; PRO1122 polypeptide; clone DNA62377-1381-1; UNQ561;
KW PCR primer; probe; antagonist; degenerative cartilaginous disorder;
KW agonist; virtual DNA fragment; DNA49665; ss.
XX
OS Synthetic.
XX
PN WO9960127-A2.
XX
PD 25-NOV-1999.
XX
PF 14-MAY-1999; 99WO-US10733.
XX
PR 15-MAY-1998; 98US-0085579.
PR 23-DEC-1998; 98US-0113621.
XX
PA (GETH) GENENTECH INC.
XX
PI Chen J, Filvaroff E, Goddard A, Gurney AL, Li H, Wood WI;
XX

```
DR   WPI; 2000-116314/10.
XX
PT   New polypeptides designated PRO1031 and PRO1122 used to treat a
PT   degenerative cartilaginous disorder -
XX
PS   Example 2; Fig 6; 141pp; English.
XX
CC   The present sequence is the virtual DNA fragment DNA49665, used for the
CC   isolation of the clone DNA62377-1381-1, encoding human PRO1122
CC   polypeptide, IL-17C. This sequence is used to generate oligonucleotides
CC   for use as PCR primers to identify the cDNA library that contains the
CC   sequence of interest and for use as probes for isolation of clone with
CC   the full-length coding sequence of PRO1122. The PRO1122 polypeptide,
CC   agonist or antagonist, can be used to diagnose and treat a degenerative
CC   cartilaginous disorder.
XX
SQ   Sequence 230 BP; 32 A; 88 C; 70 G; 40 T; 0 other;

  Query Match             32.6%;  Score 230;  DB 21;  Length 230;
  Best Local Similarity   100.0%;  Pred. No. 1e-35;
  Matches  230;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 cacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgcagaggctg 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 cacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgcagaggctg 60

Qy      73 tatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctgctccagag 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 tatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctgctccagag 120

Qy     133 cctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctccccacacctgg 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 cctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctccccacacctgg 180

Qy     193 ggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacct 242
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacct 230


RESULT   6
Z36856
ID   Z36856 standard; DNA; 332 BP.
XX
AC   Z36856;
XX
DT   13-MAR-2000  (first entry)
XX
DE   cDNA clone HTGED19RB related to human interleukin-22 CDNA sequence.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma; ss.
XX
OS   Homo sapiens.
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Disclosure; Page 168; 170pp; English.
XX
CC   The present sequence represents a cDNA clone that is related to cDNA
CC   encoding human interleukin-22 (IL-22) protein. The specification also
CC   describes IL-21 polynucleotides and polypeptides. The IL-21
CC   polynucleotide was isolated from a cDNA library of apoptotic T-cells.
CC   IL-21 and IL-22 may be useful in treating deficiencies or disorders of
CC   the immune system, by activating or inhibiting the proliferation,
CC   differentiation, or mobilization (chemotaxis) of immune cells, treating
CC   or detecting deficiencies or disorders of haematopoietic cells, to
CC   modulate haemostatic or thrombolytic activity, in treating or detecting
CC   autoimmune disorders, treating asthma (particularly allergic asthma)
CC   or other respiratory problems, to treat and/or prevent organ rejection
CC   or graft-versus-host disease (GVHD), to modulate inflammation, to
CC   treat or detect hyperproliferative disorders, to treat or detect
CC   infectious agents, to differentiate, proliferate and attract cells,
CC   leading to the regeneration of tissues, IL-21 and IL-22 may also
CC   increase or decrease the differentiation or proliferation of embryonic
CC   stem cells and haematopoietic lineage, may be used to modulate mammalian
CC   characteristics.
XX
SQ   Sequence 332 BP; 55 A; 112 C; 97 G; 60 T; 8 other;

  Query Match             20.0%;  Score 141.2;  DB 21;  Length 332;
  Best Local Similarity   88.9%;  Pred. No. 8.2e-19;
  Matches  176;  Conservative    0;  Mismatches   15;  Indels    7;  Gaps    2;

Qy       9 tggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgcagag 68
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tggacacgtatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgcagag 60

Qy      69 gctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctgctcc 128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 gctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctgctcc 120

Qy     129 agagcctg-ctggtgctgcgccgccggccctg------ctcccgcgacggctcggggctc 181
           |||||||| |||||||||||||||||||||||      | |    |||  |  |||   |
Db     121 agagcctgactggtgctgcgccgccggccctgactacccgcnacggactacgggggctac 180

Qy     182 cccacacctggggccttt 199
           |||||||||||||| | ||
Db     181 cccacacctggggncctt 198


RESULT   7
Q50573
ID   Q50573 standard; cDNA to mRNA; 1453 BP.
XX
AC   Q50573;
XX
DT   24-MAY-1994  (first entry)
XX
DE   Asparaginylendopeptidase clone 104.
XX
KW   Asparaginylendopeptidase; Canavalia ensiformis; seed;
KW   L-asparagine; primer; PCR; protein fragmentation;
KW   peptide synthesis; ss.
XX
OS   Canavalia ensiformis.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..1094
FT                   /*tag=  a
XX
PN   JP05276960-A.
```

```
XX
PD   26-OCT-1993.
XX
PF   07-AUG-1992;   92JP-0231602.
XX
PR   07-FEB-1992;   92JP-0056023.
XX
PA   (SYUZ/) SYUZO T.
XX
DR   WPI; 1993-373587/47.
DR   P-PSDB; R43038.
XX
PT   New gene for encoding asparaginyl endo-peptidase - comprises 8
PT   specified DNA sequences
XX
PS   Disclosure; Page 22-24; 35pp; Japanese.
XX
CC   A gene encoding asparaginylendopeptidase is claimed.
CC   8 DNA sequences are given (Q50559-66).  The enzyme is a protease
CC   derived from a seed of Canavalia ensiformis which selectively
CC   hydrolyses C-terminus amide bond of L-asparagine residue
CC   (see R43033 and R43041).
CC   THe enzyme is useful for protein fragmentation and enzymatic
CC   peptide synthesis.
CC   The primers given in Q50567-68, Q50576-77 and Q50583-90 were used in
CC   the isolation of the fragments given in Q50569-75 and Q50578-79,
CC   by PCR.
XX
SQ   Sequence 1453 BP; 480 A; 252 C; 335 G; 386 T; 0 other;


  Query Match             9.5%;  Score 67;  DB 14;  Length 1453;
  Best Local Similarity   79.8%;  Pred. No. 0.00013;
  Matches   79;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy     607 cctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaa 666
           ||| ||  | ||| || || |||||| |  || |||||  |||   |  || | ||||||
Db    1343 cctttttcataaaataagtaattaagtattagtctattaaaaaaaaaaaaaaaaaaaaaa 1402

Qy     667 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||||
Db    1403 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1441


RESULT   8
Q50575
ID   Q50575 standard; cDNA to mRNA; 1640 BP.
XX
AC   Q50575;
XX
DT   24-MAY-1994  (first entry)
XX
DE   Asparaginylendopeptidase clone ASN-1.
XX
KW   Asparaginylendopeptidase; Canavalia ensiformis; seed;
KW   L-asparagine; primer; PCR; protein fragmentation;
KW   peptide synthesis; ss.
XX
OS   Canavalia ensiformis.
XX
PN   JP05276960-A.
XX
PD   26-OCT-1993.
XX
PF   07-AUG-1992;   92JP-0231602.
XX
PR   07-FEB-1992;   92JP-0056023.
XX
PA   (SYUZ/) SYUZO T.
XX
DR   WPI; 1993-373587/47.
XX
PT   New gene for encoding asparaginyl endo-peptidase - comprises 8
PT   specified DNA sequences
XX
PS   Disclosure; Page 26; 35pp; Japanese.
XX
CC   A gene encoding asparaginylendopeptidase is claimed.
CC   8 DNA sequences are given (Q50559-66).  The enzyme is a protease
CC   derived from a seed of Canavalia ensiformis which selectively
CC   hydrolyses C-terminus amide bond of L-asparagine residue
CC   (see R43033 and R43041).
CC   THe enzyme is useful for protein fragmentation and enzymatic
CC   peptide synthesis.
CC   The primers given in Q50567-68, Q50576-77 and Q50583-90 were used in
CC   the isolation of the fragments given in Q50569-75 and Q50578-79,
CC   by PCR.
XX
SQ   Sequence 1640 BP; 533 A; 288 C; 383 G; 436 T; 0 other;


  Query Match             9.5%;  Score 67;  DB 14;  Length 1640;
  Best Local Similarity   79.8%;  Pred. No. 0.00013;
  Matches   79;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy     607 cctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaa 666
           ||| ||  | ||| || || |||||| |  || |||||  |||   |  || | ||||||
Db    1530 cctttttcataaaataagtaattaagtattagtctattaaaaaaaaaaaaaaaaaaaaaa 1589

Qy     667 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||||
Db    1590 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1628


RESULT   9
Q50579
ID   Q50579 standard; cDNA to mRNA; 1910 BP.
XX
AC   Q50579;
XX
DT   24-MAY-1994  (first entry)
XX
DE   Asparaginylendopeptidase ASN.
XX
KW   Asparaginylendopeptidase; Canavalia ensiformis; seed;
KW   L-asparagine; primer; PCR; protein fragmentation;
KW   peptide synthesis; ss.
XX
OS   Canavalia ensiformis.
XX
FH   Key             Location/Qualifiers
FT   CDS             229..1551
FT                   /*tag=  a
XX
PN   JP05276960-A.
XX
PD   26-OCT-1993.
XX
PF   07-AUG-1992;   92JP-0231602.
XX
PR   07-FEB-1992;   92JP-0056023.
XX
PA   (SYUZ/) SYUZO T.
XX
DR   WPI; 1993-373587/47.
DR   P-PSDB; R43040.
XX
PT   New gene for encoding asparaginyl endo-peptidase - comprises 8
PT   specified DNA sequences
XX
PS   Disclosure; Page 27-29; 35pp; Japanese.
XX
CC   A gene encoding asparaginylendopeptidase is claimed.
CC   8 DNA sequences are given (Q50559-66).  The enzyme is a protease
CC   derived from a seed of Canavalia ensiformis which selectively
CC   hydrolyses C-terminus amide bond of L-asparagine residue
```

```
CC   (see R43033 and R43041).
CC   THe enzyme is useful for protein fragmentation and enzymatic
CC   peptide synthesis.
CC   The primers given in Q50567-68, Q50576-77 and Q50583-90 were used in
CC   the isolation of the fragments given in Q50569-75 and Q50578-79,
CC   by PCR.
XX
SQ   Sequence 1910 BP; 591 A; 339 C; 464 G; 516 T; 0 other;


  Query Match             9.5%;  Score 67;  DB 14;  Length 1910;
  Best Local Similarity   79.8%;  Pred. No. 0.00014;
  Matches   79;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy     607 cctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaa 666
           ||| ||  | ||| || || |||||| |  || |||||  |||   |  || | ||||||
Db    1800 ccttttcataaaataagtaattaagtattagtctattaaaaaaaaaaaaaaaaaaaaaa 1859

Qy     667 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||||
Db    1860 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1898


RESULT  10
A23428
ID   A23428 standard; cDNA; 2010 BP.
XX
AC   A23428;
XX
DT   19-JUN-2000  (first entry)
XX
DE   cDNA encoding human secreted protein vg2_1, SEQ ID NO:11.
XX
KW   Human; secreted protein; cancer; tumour; cardiovascular disorder;
KW   blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
KW   infection; fungal; bacterial; viral; HIV; allergy; arthritis;
KW   neurodegenerative disease; asthma; contraceptive; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             124..1644
FT                   /*tag=  a
FT                   /product= "Human secreted protein vg2_1"
XX
PN   WO200011015-A1.
XX
PD   02-MAR-2000.
XX
PF   24-AUG-1999;   99WO-US19351.
XX
PR   24-AUG-1998;   98US-0097638.
PR   24-AUG-1998;   98US-0097659.
PR   09-SEP-1998;   98US-0099618.
PR   28-SEP-1998;   98US-0102092.
PR   25-NOV-1998;   98US-0109978.
PR   23-DEC-1998;   98US-0113645.
PR   23-DEC-1998;   98US-0113646.
PR   23-AUG-1999;   99US-0379246.
XX
PA   (ALPH-) ALPHAGENE INC.
XX
PI   Valenzuela D,  Yuan O,  Hoffman H,  Hall J,  Rapiejko P;
XX
DR   WPI; 2000-224657/19.
DR   P-PSDB; Y94986.
XX
PT   New secreted or transmembrane proteins and polynucleotides encoding
PT   them, useful for treating neurodegenerative disorders, autoimmune
PT   diseases and cancer -
XX
PS   Claim 20; Page 268-269; 357pp; English.
XX
CC   The invention relates to 40 human secreted proteins (Y94981-Y95020),
CC   and cDNA sequences encoding them (A23423-A23462). The secreted proteins
CC   of the invention include those that are thought to be only partially
CC   secreted, i.e., transmembrane proteins. The proteins of the invention may
CC   exhibit one or more activities selected from the following: cytokine
CC   activity; cell proliferation; differentiation; immune modulation;
CC   haematopoiesis regulation; tissue growth activity; activin/inhibin
CC   activity; chemotactic/chemokinetic activity; haemostatic and
CC   thrombolytic activity; anti-inflammatory activity; and tumour inhibition
CC   activity. The proteins may be administered to patients as vaccines, and
CC   the nucleotides may be used as part of a gene therapy regime. Diseases or
CC   conditions that may be treated using the proteins or nucleotides of the
CC   invention include autoimmune diseases; genetic disorders; haemophilia;
CC   cardiovascular diseases; cancer; bacterial, fungal and viral infections,
CC   especially HIV; multiple sclerosis; rheumatoid arthritis; pulmonary
CC   inflammation; Guillain-Barre syndrome; insulin dependent diabetes
CC   mellitus; and allergic reactions such as asthma and anaemia. They may
CC   also be used for treating wounds, burns, ulcers, osteoporosis,
CC   osteoarthritis, periodontal diseases, Alzheimer's disease, Parkinson's
CC   disease, Huntington's disease and amyotrophic lateral sclerosis (ALS).
CC   Proteins with activin/inhibin activity may additionally be useful as
CC   contraceptives. Nucleic acid sequences of the invention may be used in
CC   chromosome mapping, and as a source of diagnostic primers and probes.
CC   The present sequence represents cDNA encoding one of the 40 proteins of
CC   the invention.
XX
SQ   Sequence 2010 BP; 421 A; 661 C; 592 G; 336 T; 0 other;


  Query Match             9.3%;  Score 65.8;  DB 21;  Length 2010;
  Best Local Similarity   60.2%;  Pred. No. 0.00023;
  Matches  109;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     525 tgtcctgctcccggcttcccttaccctatcactggcctcaggccccgcaggctgcctct 584
           | |||   ||| | |  |||| | ||   |     ||   |  |||  ||||  ||| ||
Db    1808 tttccgagtccagccccccctgagccccacctcctccagggagccctccaggaagcccct 1867

Qy     585 tcccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtatta 644
           ||||   ||||| |     |  || |      ||| || |  |    |  |     | |
Db    1868 tccctgactcctggcccacaggccaggcctaaaaaaaactcgtggcttaaaaaaaaaaaa 1927

Qy     645 ttaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 704
             |||   |  || | ||||||||||||||||||||||||||||||||||||||||||
Db    1928 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1987

Qy     705 a 705
           |
Db    1988 a 1988


RESULT  11
X16675
ID   X16675 standard; DNA; 1554 BP.
XX
AC   X16675;
XX
DT   29-APR-1999  (first entry)
XX
DE   Xenopus WA545 protein encoding DNA.
XX
KW   Xenopus; WA545 protein; TGF-beta; transforming growth factor beta;
KW   proliferation; differentiation; mesodermal tissue; neural; muscle;
KW   bone; cartilage; connective tissue; wound healing; gene therapy; ss.
XX
OS   Xenopus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             55..1119
FT                   /*tag=  a
FT                   /product= "WA545 protein"
FT   sig_peptide     55..774
FT                   /*tag=  b
FT   mat_peptide     775..1116
```

```
FT                  /*tag=  c
XX
PN   WO9902678-A1.
XX
PD   21-JAN-1999.
XX
PF   24-APR-1998;   98WO-US08334.
XX
PR   10-JUL-1997;   97US-0890918.
XX
PA   (GEMY ) GENETICS INST INC.
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Lavallie ER,  Racie LA,  Sive H,  Sun B;
XX
DR   WPI; 1999-120879/10.
DR   P-PSDB; W94656.
XX
PT   New isolated Xenopus WA545 DNA - which is a member of the
PT   transforming growth factor-beta superfamily, used to induce tissue
PT   formation and in wound healing
XX
PS   Claim 1; Page 51-52; 73pp; English.
XX
CC   The present sequence encodes the Xenopus WA545 protein. WA545 proteins
CC   have the ability to induce, enhance and/or inhibit the formation,
CC   growth, proliferation, differentiation, maintenance of mesodermal
CC   tissue, including neurons and/or related neural cells and tissues such
CC   as brain cells, Schwann cells, glial cells and astrocytes, as well as
CC   muscle cells and tissues. They can be used for treating bone, cartilage,
CC   muscle, nerve, epidermis and/or other connective tissue defects, as well
CC   as periodontal disease and healing of various epidermis, nerve,
CC   including spinal cord, muscle, including cardiac, striated or smooth
CC   muscle, and other tissues and wounds, and other organs such as liver,
CC   pancreas, spleen brain, lung, cardiac and kidney tissue. They can also
CC   be used to treat or prevent such conditions as osteoarthritis,
CC   osteoporosis, and other abnormalities of bone, cartilage, muscle, nerve,
CC   epidermis or other connective tissue, organs such as liver, pancreas,
CC   spleen, lung, cardiac and kidney and other tissues. They can also be
CC   used for wound healing, reduction of fibrosis and reduction of scar
CC   tissue formation. They can also be used to induce bone and/or cartilage
CC   or other connective tissue formation, and in wound healing and tissue
CC   repair and for augmenting the activity of bone morphogenetic proteins.
CC   They can also be used as a dietary or nutrient supplement. The products
CC   can also be used for treating cell populations, such as embryonic cells
CC   or stem cell populations to enhance or enrich the growth and/or
CC   differentiation of the cells and for implantation and for gene therapy
CC   applications.
XX
SQ   Sequence 1554 BP; 503 A; 340 C; 309 G; 402 T; 0 other;


  Query Match             8.7%;  Score 61.4;  DB 20;  Length 1554;
  Best Local Similarity   72.1%;  Pred. No. 0.0015;
  Matches   80;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy     595 cttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgat 654
           |||| ||||||  ||| || ||  | |   || | | |  ||    |  |  |||   |
Db    1437 cttgcaagtacagctgattgttggaaataaatgtgatttaacccaaaaaaaaaaaaaaaa 1496

Qy     655 gaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           || | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1497 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1547


RESULT  12
T72173
ID   T72173 standard; cDNA to mRNA; 882 BP.
XX
AC   T72173;
XX
DT   25-FEB-1998  (first entry)
XX
DE   Alzheimer's disease DNA sequence from plasmid pGCS1180.
XX
KW   Human; brain; Alzheimer's disease; diagnosis; antibody; expression; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             97..198
FT                   /*tag=  a
FT                   /note= "Protein encoded by DNA from the human brain which
FT                          has different expression from the normal protein
FT                          when in an Alzheimer's disease patient"
XX
PN   WO9721807-A1.
XX
PD   19-JUN-1997.
XX
PF   12-DEC-1996;   96WO-JP03630.
XX
PR   12-DEC-1995;   95JP-0322745.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Hashida H,  Kuga T,  Nakagawa S,  Sakaki Y,  Zhao N;
XX
DR   WPI; 1997-332779/30.
DR   P-PSDB; W19950.
XX
PT   DNA sequences with varied expression in normal and Alzheimer
PT   patients - used for diagnosis and treatment of Alzheimer's disease
XX
PS   Claim 3; Page 48-49; 90pp; Japanese.
XX
CC   The present sequence represents a DNA fragment which is complementary to
CC   mRNA sequences whose level of expression in the brains of Alzheimer's
CC   disease patients varies from normal individuals. The present sequence
CC   represents a specifically claimed DNA fragment in which the expression
CC   is augmented to at least twice that in normal cases. The sequence, and
CC   antibodies to the proteins encoded by the sequence, may be used in the
CC   diagnosis, treatment and investigation of Alzheimer's disease.
XX
SQ   Sequence 882 BP; 306 A; 134 C; 142 G; 300 T; 0 other;


  Query Match             8.7%;  Score 61.2;  DB 18;  Length 882;
  Best Local Similarity   76.5%;  Pred. No. 0.0015;
  Matches   75;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy     608 ctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaa 667
           || |   || |  ||||  |||  |  |    |||||| ||| |||  || | |||||||
Db     785 ctttcaatttataaatttgttatcttaataaatattataaaaatgaaaaaaaaaaaaaaa 844

Qy     668 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           ||||||||||||||||||||||||||||||||||||||||
Db     845 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 882


RESULT  13
V61487
ID   V61487 standard; cDNA; 4237 BP.
XX
AC   V61487;
XX
DT   11-JAN-1999  (first entry)
XX
DE   Human secreted protein fe366_1 cDNA.
XX
KW   Secreted protein; human; fe366_1; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3746..4030
```

```
FT                  /*tag=  a
XX
PN   WO9841539-A2.
XX
PD   24-SEP-1998.
XX
PF   19-MAR-1998;   98WO-US05474.
XX
PR   18-MAR-1998;   98US-0040963.
PR   19-MAR-1997;   97US-0820493.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Agostino MJ,  Jacobs K,  LaVallie ER,  McCoy JM,  Merberg D;
PI   Racie LA,  Spaulding V,  Treacy M;
XX
DR   WPI; 1998-521163/44.
DR   P-PSDB; W79096.
XX
PT   New polynucleotide(s) encoding secreted human proteins - derived
PT   from human foetal kidney, adult testes and adult or foetal brain
PT   cDNA libraries
XX
PS   Claim 36; Page 88-91; 112pp; English.
XX
CC   This full-length cDNA clone, designated fe366_1, codes for a novel
CC   secreted human protein (see W79096).  It was isolated from a human
CC   adult brain cDNA library using methods which are selective for
CC   cDNAs encoding secreted proteins, or was identified as encoding
CC   a secreted or transmembrane protein on the basis of computer
CC   analysis of the encoding protein.  The nucleotide sequence shows
CC   homology to some database sequences, and may contain a CAA repeat
CC   and/or Alu repetitive element.  The invention provides cDNA clones
CC   (see V61477-87) from human foetal kidney, adult testis, and adult
CC   or foetal brain cDNA libraries that code for secreted proteins
CC   (see W79087-97).  These clones are deposited as ATCC 98364.  The
CC   polynucleotides and proteins are predicted to have useful
CC   biological activities which would make them suitable for treating,
CC   preventing or ameliorating medical conditions in humans and
CC   animals, although no supporting data is given.  Suggested
CC   activities include nutritional, immune stimulating (e.g. as
CC   vaccines) or suppressing, haematopoiesis regulating, tissue growth,
CC   activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC   thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour
CC   invasion suppressor and tumour inhibition activities.  The
CC   polynucleotides are also stated to be useful for gene therapy.
XX
SQ   Sequence 4237 BP; 1330 A; 778 C; 784 G; 1338 T; 7 other;


  Query Match             8.7%;  Score 61.2;  DB 19;  Length 4237;
  Best Local Similarity   75.0%;  Pred. No. 0.0019;
  Matches   75;  Conservative    1;  Mismatches   24;  Indels    0;  Gaps    0;

Qy     606 ccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaa 665
           || | ||     || || | |||||||  |  ||   |   ||| |||  || |:||||
Db    2380 cctttttctcccaaaaaatctttaagtaaagatgatctgggaaaatgaaaaaaasaaaaa 2439

Qy     666 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||||||
Db    2440 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa.2479


RESULT  14
V59803
ID   V59803 standard; DNA; 1181 BP.
XX
AC   V59803;
XX
DT   19-JAN-1999  (first entry)
XX
DE   Human secreted protein gene 176 clone HFTBR48.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9839448-A2.
XX
PD   11-SEP-1998.
XX
PF   06-MAR-1998;   98WO-US04493.
XX
PR   02-OCT-1997;   97US-0061060.
PR   07-MAR-1997;   97US-0038621.
PR   07-MAR-1997;   97US-0040161.
PR   07-MAR-1997;   97US-0040162.
PR   07-MAR-1997;   97US-0040163.
PR   07-MAR-1997;   97US-0040333.
PR   07-MAR-1997;   97US-0040334.
PR   07-MAR-1997;   97US-0040336.
PR   07-MAR-1997;   97US-0040626.
PR   11-APR-1997;   97US-0043311.
PR   11-APR-1997;   97US-0043312.
PR   11-APR-1997;   97US-0043313.
PR   11-APR-1997;   97US-0043314.
PR   11-APR-1997;   97US-0043568.
PR   11-APR-1997;   97US-0043569.
PR   11-APR-1997;   97US-0043576.
PR   11-APR-1997;   97US-0043578.
PR   11-APR-1997;   97US-0043580.
PR   11-APR-1997;   97US-0043669.
PR   11-APR-1997;   97US-0043670.
PR   11-APR-1997;   97US-0043671.
PR   11-APR-1997;   97US-0043672.
PR   11-APR-1997;   97US-0043674.
PR   23-MAY-1997;   97US-0047492.
PR   23-MAY-1997;   97US-0047500.
PR   23-MAY-1997;   97US-0047501.
PR   23-MAY-1997;   97US-0047502.
PR   23-MAY-1997;   97US-0047503.
PR   23-MAY-1997;   97US-0047581.
PR   23-MAY-1997;   97US-0047582.
PR   23-MAY-1997;   97US-0047583.
PR   23-MAY-1997;   97US-0047584.
PR   23-MAY-1997;   97US-0047585.
PR   23-MAY-1997;   97US-0047586.
PR   23-MAY-1997;   97US-0047587.
PR   23-MAY-1997;   97US-0047588.
PR   23-MAY-1997;   97US-0047589.
PR   23-MAY-1997;   97US-0047590.
PR   23-MAY-1997;   97US-0047592.
PR   23-MAY-1997;   97US-0047593.
PR   23-MAY-1997;   97US-0047594.
PR   23-MAY-1997;   97US-0047595.
PR   23-MAY-1997;   97US-0047596.
PR   23-MAY-1997;   97US-0047597.
PR   23-MAY-1997;   97US-0047598.
PR   23-MAY-1997;   97US-0047599.
PR   23-MAY-1997;   97US-0047600.
PR   23-MAY-1997;   97US-0047601.
PR   23-MAY-1997;   97US-0047612.
PR   23-MAY-1997;   97US-0047613.
PR   23-MAY-1997;   97US-0047614.
PR   23-MAY-1997;   97US-0047615.
PR   23-MAY-1997;   97US-0047617.
PR   23-MAY-1997;   97US-0047618.
PR   23-MAY-1997;   97US-0047632.
PR   23-MAY-1997;   97US-0047633.
```

PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057669.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
DR WPI; 1998-506364/43.
DR P-PSDB; W75017.
XX
PT New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 1; Page 534-535; 721pp; English.
XX
CC This sequence represents a nucleic acid molecule designated Gene 176 from
CC the human cDNA clone HFTBR48 (deposited as clone ATCC 97904 and ATCC
CC 209050) which encodes a secreted human protein. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. V59502) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 186 polynucleotides, based on
CC which tissues they are most highly expressed in (see V59511 for described
CC uses).
XX
SQ Sequence 1181 BP; 357 A; 228 C; 295 G; 298 T; 3 other;

Query Match 8.7%; Score 61; DB 19; Length 1181;
Best Local Similarity 64.5%; Pred. No. 0.0017;
Matches 91; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 565 ggccccgcaggctgcctcttcccaacctccttggaagtacccctgtttcttaaacaatt 624
Db 1033 gtccatcccttggtacatatatgcatttgcttttaaaccatttcttttgtttaaataaat 1092

Qy 625 atttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaa 684
Db 1093 aaataagtaaataaagctagttctattgaaatgcaaaaaaaaaaaaaaaaaaaaaaaaaa 1152

Qy 685 aaaaaaaaaaaaaaaaaaaa 705
Db 1153 aaaaaaaaaaaaaaaaaaaa 1173

RESULT 15
V59686
ID V59686 standard; DNA; 1212 BP.
XX
AC V59686;
XX
DT 19-JAN-1999 (first entry)
XX
DE Human secreted protein gene 176 clone HFTBR48.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9839448-A2.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US04493.
XX
PR 02-OCT-1997; 97US-0061060.
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.

PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057669.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
DR WPI; 1998-506364/43.
DR P-PSDB; W74903.
XX
PT New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 1; Page 415-416; 721pp; English.
XX
CC This sequence represents a nucleic acid molecule designated Gene 176 from
CC the human cDNA clone HFTBR48 (deposited as clone ATCC 97904 and ATCC
CC 209050) which encodes a secreted human protein. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. V59502) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: V59511-V59812; amino acid sequences W74731-W75026)' which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 186 polynucleotides, based on
CC which tissues they are most highly expressed in (see V59511 for described
CC uses).
XX
SQ Sequence 1212 BP; 363 A; 241 C; 307 G; 300 T; 1 other;

Query Match 8.7%; Score 61; DB 19; Length 1212;
Best Local Similarity 64.5%; Pred. No. 0.0017;
Matches 91; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```
Qy    565 ggcccccgcaggctgcctcttcccaacctccttggaagtacccctgtttcttaaacaatt 624
          | ||  | |  | | | | | |   ||    | ||  ||  |   || ||  ||||| || |
Db   1064 gtccatcccttggtacatatatgcatttgcttttaaaccatttcttttgtttaaataaat 1123

Qy    625 atttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaa 684
          |  |||||  |      |  ||  | |||    || ||||||||||||||||||||||||
Db   1124 aaataagtaaataaagctagttctattgaaatgcaaaaaaaaaaaaaaaaaaaaaaaaa 1183

Qy    685 aaaaaaaaaaaaaaaaaaaa 705
          ||||||||||||||||||||
Db   1184 aaaaaaaaaaaaaaaaaaaa 1204
```

Search completed: February 24, 2001, 21:51:30
Job time: 7082 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 18:18:12 ; Search time 2862.16 Seconds
(without alignments)
1260.591 Million cell updates/sec

Title: US-09-320-713-1
Perfect score: 705
Sequence: 1 ggcacgagtggacacggatg.....aaaaaaaaaaaaaaaaaaaa 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_ph:*
6: gb_pl1:*
7: gb_pl2:*
8: gb_pr1:*
9: gb_pr2:*
10: gb_pr3:*
11: em_fun:*
12: em_hum1:*
13: em_hum2:*
14: em_in:*
15: em_om:*
16: em_or:*
17: em_ov:*
18: em_pat:*
19: em_ph:*
20: em_pl:*
21: em_ro:*
22: em_sts:*
23: em_sy:*
24: em_un:*
25: em_vi:*
26: gb_htg1:*
27: gb_htg2:*
28: gb_in1:*
29: gb_in2:*
30: em_ba1:*
31: em_ba2:*
32: em_hum3:*
33: em_hum4:*
34: gb_pr4:*
35: gb_htg3:*
36: gb_htg4:*
37: gb_htg5:*
38: gb_htg6:*
39: gb_htg7:*
40: em_htg1:*
41: em_htg2:*
42: em_htg3:*
43: em_hum5:*
44: gb_pl3:*
45: gb_pr5:*
46: gb_htg8:*
47: gb_htg9:*
48: gb_htg10:*
49: gb_htg11:*
50: gb_htg12:*
51: gb_htg13:*
52: gb_htg14:*
53: gb_in3:*
54: gb_htg15:*
55: gb_htg16:*
56: gb_htg17:*
57: em_htg4:*
58: em_htg5:*
59: em_htg6:*
60: em_htg7:*
61: em_hum6:*
62: gb_htg18:*
63: gb_htg19:*
64: gb_ba3:*
65: em_htg8:*
66: em_htg9:*
67: em_htg10:*
68: gb_pr6:*
69: gb_pr7:*
70: gb_htg20:*
71: gb_htg21:*
72: gb_htg22:*
73: gb_htg23:*
74: gb_ro:*
75: gb_sts1:*
76: gb_sts2:*
77: gb_sy:*
78: gb_un:*
79: gb_vi1:*
80: gb_vi2:*
81: gb_pat1:*
82: gb_pat2:*
83: em_htg0:*
84: gb_htg24:*
85: gb_pr8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	650.4	92.3	157090	47	AC022554	AC022554 Homo sapi
2	640.4	90.8	1047	34	AF152099	AF152099 Homo sapi
3	64.2	9.1	1315	74	AF082526	AF082526 Mus muscu
4	61.8	8.8	989	7	CPCRM7	Y08243 C.pteroides
5	61.4	8.7	1554	4	AF065135	AF065135 Xenopus l
6	61.2	8.7	1221	69	HSM801506	AL133637 Homo sapi
7	60.8	8.6	3469	69	HSM800438	AL050138 Homo sapi
8	60.8	8.6	6410	34	AF230496	AF230496 Homo sapi
9	60.4	8.6	1478	34	AK025967	AK025967 Homo sapi
10	60.4	8.6	1678	69	HSM802263	AL137530 Homo sapi
11	60.2	8.5	2358	3	AF254119	AF254119 Ovis arie
12	60	8.5	1419	7	CCCHSMR	Z67988 C.chinensis
13	60	8.5	3236	69	HSM801246	AL122093 Homo sapi
14	59.8	8.5	1809	7	ATPUR2	X74766 Arabidopsis
15	59.8	8.5	4173	69	HSM801960	AL137292 Homo sapi
16	59.6	8.5	7218	81	I66494	I66494 Sequence 14
17	59.4	8.4	1523	10	AF116614	AF116614 Homo sapi
c 18	59.4	8.4	49473	52	AC061987	AC061987 Homo sapi
19	59.2	8.4	2156	34	AK025435	AK025435 Homo sapi
20	59	8.4	1560	34	AK026057	AK026057 Homo sapi
21	59	8.4	1739	44	PCU42442	U42442 Pneumocysti

22	59	8.4	1826	7	AF196975	AF196975 Pneumocys
23	58.8	8.3	1056	74	MUSUNKNB	L04849 Mouse (clon
24	58.8	8.3	1405	69	HSM800974	AL117463 Homo sapi
25	58.8	8.3	2341	10	AF119857	AF119857 Homo sapi
26	58.8	8.3	2345	74	MMU87620	U87620 Mus musculu
27	58.8	8.3	4212	69	HSM800772	AL110269 Homo sapi
28	58.6	8.3	1740	10	AF069506	AF069506 Homo sapi
29	58.6	8.3	1845	69	HSM800467	AL050172 Homo sapi
30	58.6	8.3	110000	84	PFMAL13P2_0	AL049185 Plasmodiu
31	58.4	8.3	617	85	IROEST087	AL357195 Homo sapi
32	58.4	8.3	2079	8	AB047878	AB047878 Macaca fa
33	58.4	8.3	2665	34	AF260566	AF260566 Homo sapi
34	58.2	8.3	1602	81	AR000496	AR000496 Sequence
35	58.2	8.3	1699	69	HSU39656	U39656 Human MAP k
36	58.2	8.3	1798	74	AF246457	AF246457 Rattus no
37	58.2	8.3	1838	74	AF246458	AF246458 Rattus no
38	58.2	8.3	1897	81	I52013	I52013 Sequence 1
39	58	8.2	1288	53	TBBDPMSYN	Z54162 T.brucei br
40	58	8.2	2344	34	AK024855	AK024855 Homo sapi
41	58	8.2	2827	69	HSM801964	AL137660 Homo sapi
42	58	8.2	4632	69	HSU72621	U72621 Homo sapien
43	57.8	8.2	522	34	AK026629	AK026629 Homo sapi
44	57.8	8.2	721	34	AK026526	AK026526 Homo sapi
45	57.8	8.2	1162	7	AF275315	AF275315 Lotus jap

ALIGNMENTS

```
RESULT   1
AC022554
LOCUS       AC022554   157090 bp    DNA             HTG       22-APR-2000
DEFINITION  Homo sapiens clone RP11-21B21, WORKING DRAFT SEQUENCE, 35 unordered
            pieces.
ACCESSION   AC022554
VERSION     AC022554.2  GI:7637249
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 157090)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome, clone RP11-21B21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157090)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 22, 2000 this sequence version replaced gi:6910831.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2245
                Center clone name: 21_B_21
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 138238 bases at least Q40
                Consensus quality: 147017 bases at least Q30
                Consensus quality: 150850 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 153690; sum-of-contigs
                Quality coverage: 3.1 in Q20 bases; agarose-fp
                Quality coverage: 3.5 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 35 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1661: contig of 1661 bp in length
            *     1662  1761: gap of      100 bp
            *     1762     3078: contig of 1317 bp in length
            *     3079  3178: gap of      100 bp
            *     3179     4774: contig of 1596 bp in length
            *     4775  4874: gap of      100 bp
            *     4875     7439: contig of 2565 bp in length
            *     7440  7539: gap of      100 bp
            *     7540     8885: contig of 1346 bp in length
            *     8886  8985: gap of      100 bp
            *     8986    11624: contig of 2639 bp in length
            *    11625 11724: gap of      100 bp
            *    11725    14597: contig of 2873 bp in length
            *    14598 14697: gap of      100 bp
            *    14698    16860: contig of 2163 bp in length
            *    16861 16960: gap of      100 bp
            *    16961    19655: contig of 2695 bp in length
            *    19656 19755: gap of      100 bp
            *    19756    21926: contig of 2171 bp in length
            *    21927 22026: gap of      100 bp
            *    22027    23792: contig of 1766 bp in length
            *    23793 23892: gap of      100 bp
            *    23893    27939: contig of 4047 bp in length
            *    27940 28039: gap of      100 bp
            *    28040    30269: contig of 2230 bp in length
            *    30270 30369: gap of      100 bp
            *    30370    33004: contig of 2635 bp in length
            *    33005 33104: gap of      100 bp
            *    33105    36260: contig of 3156 bp in length
            *    36261 36360: gap of      100 bp
            *    36361    39762: contig of 3402 bp in length
            *    39763 39862: gap of      100 bp
            *    39863    43798: contig of 3936 bp in length
            *    43799 43898: gap of      100 bp
            *    43899    48559: contig of 4661 bp in length
            *    48560 48659: gap of      100 bp
            *    48660    52026: contig of 3367 bp in length
            *    52027 52126: gap of      100 bp
            *    52127    55169: contig of 3043 bp in length
            *    55170 55269: gap of      100 bp
            *    55270    58796: contig of 3527 bp in length
            *    58797 58896: gap of      100 bp
            *    58897    62413: contig of 3517 bp in length
            *    62414 62513: gap of      100 bp
            *    62514    67190: contig of 4677 bp in length
            *    67191 67290: gap of      100 bp
            *    67291    71331: contig of 4041 bp in length
            *    71332 71431: gap of      100 bp
```

```
            *    71432    75280: contig of 3849 bp in length
            *    75281 75380: gap of      100 bp
            *    75381    80896: contig of 5516 bp in length
            *    80897 80996: gap of      100 bp
            *    80997    87850: contig of 6854 bp in length
            *    87851 87950: gap of      100 bp
            *    87951    93903: contig of 5953 bp in length
            *    93904 94003: gap of      100 bp
            *    94004   101488: contig of 7485 bp in length
            *   101489 101588: gap of      100 bp
            *   101589   109837: contig of 8249 bp in length
            *   109838 109937: gap of      100 bp
            *   109938   116942: contig of 7005 bp in length
            *   116943 117042: gap of      100 bp
            *   117043   124191: contig of 7149 bp in length
            *   124192 124291: gap of      100 bp
            *   124292   133924: contig of 9633 bp in length
            *   133925 134024: gap of      100 bp
            *   134025   144057: contig of 10033 bp in length
            *   144058 144157: gap of      100 bp
            *   144158   157090: contig of 12933 bp in length.
FEATURES             Location/Qualifiers
     source          1. .157090
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-21B21"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1661
                     /note="assembly_fragment"
     misc_feature    1762. .3078
                     /note="assembly_fragment"
     misc_feature    3179. .4774
                     /note="assembly_fragment"
     misc_feature    4875. .7439
                     /note="assembly_fragment"
     misc_feature    7540. .8885
                     /note="assembly_fragment"
     misc_feature    8986. .11624
                     /note="assembly_fragment"
     misc_feature    11725. .14597
                     /note="assembly_fragment"
     misc_feature    14698. .16860
                     /note="assembly_fragment"
     misc_feature    16961. .19655
                     /note="assembly_fragment"
     misc_feature    19756. .21926
                     /note="assembly_fragment"
     misc_feature    22027. .23792
                     /note="assembly_fragment"
     misc_feature    23893. .27939
                     /note="assembly_fragment"
     misc_feature    28040. .30269
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    30370. .33004
                     /note="assembly_fragment"
     misc_feature    33105. .36260
                     /note="assembly_fragment"
     misc_feature    36361. .39762
                     /note="assembly_fragment"
     misc_feature    39863. .43798
                     /note="assembly_fragment"
     misc_feature    43899. .48559
                     /note="assembly_fragment"
     misc_feature    48660. .52026
                     /note="assembly_fragment"
     misc_feature    52127. .55169
                     /note="assembly_fragment"
     misc_feature    55270. .58796
                     /note="assembly_fragment"
     misc_feature    58897. .62413
                     /note="assembly_fragment"
     misc_feature    62514. .67190
                     /note="assembly_fragment"
     misc_feature    67291. .71331
                     /note="assembly_fragment"
     misc_feature    71432. .75280
                     /note="assembly_fragment"
     misc_feature    75381. .80896
                     /note="assembly_fragment"
     misc_feature    80997. .87850
                     /note="assembly_fragment"
     misc_feature    87951. .93903
                     /note="assembly_fragment"

  Query Match          92.3%;  Score 650.4;  DB 47;  Length 157090;
  Best Local Similarity   99.2%;  Pred. No. 2.4e-114;
  Matches  651;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      6 gagtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgca 65
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125152 GTGTGGACACGGATGAGGACCGCTATCCACAGAAGCTGGCCTTCGCCGAGTGCCTGTGCA 125211

Qy     66 gaggctgtatcgatgcacggacgggccgcgagacagctgggctcaactccgtgcggctgc 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125212 GAGGCTGTATCGATGCACGGACGGGCCGCGAGACAGCTGCGCTCAACTCCGTGCGGCTGC 125271

Qy    126 tccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctcccca 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125272 TCCAGAGCCTGCTGGTGCTGCGCCGCCGGCCCTGCTCCCGCGACGGCTCGGGGCTCCCCA 125331

Qy    186 cacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacctgcg 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125332 CACCTGGGGCCTTTGCCTTCCACACCGAGTTCATCCACGTCCCCGTCGGCTGCACCTGCG 125391

Qy    246 tgctgccccgttcagtgtgaccgccaaggccgtggggcccttagactggacacgtgtgct 305
          ||||||||||||||||||||||| ||||||||||||| |||||||||||||||||||||
Db 125392 TGCTGCCCCGTTCAGTGTGACCGCCGAGGCCGTGGGGCCCCTAGACTGGACACGTGTGCT 125451

Qy    306 ccccagagggcacccccctatttatgtgtatttattgttatttatatgcctcccccaacac 365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125452 CCCCAGAGGGCACCCCCTATTTATGTGTATTTATTGTTATTTATATGCCTCCCCCAACAC 125511

Qy    366 taccccttggggtctgggcattccccgtgtctggaggacagccccccactgttctcctcat 425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125512 TACCCTTGGGGTCTGGGCATTCCCCGTGTCTGGAGGACAGCCCCCCACTGTTCTCCTCAT 125571

Qy    426 ctccagcctcagtagttgggggtwgaaggagctcagcacctcttccagcccttaaagctg 485
          |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db 125572 CTCCAGCCTCAGTAGTTGGGGGTAGAAGGAGCTCAGCACCTCTTCCAGCCCTTAAAGCTG 125631

Qy    486 cagaaaaggtgtcacacggctgcctgtaccttggytccctgtcctgctcccggcttccct 545
          ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db 125632 CAGAAAAGGTGTCACACGGCTGCCTGTACCTTGGCTCCCTGTCCTGCTCCCGGCTTCCCT 125691

Qy    546 taccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttggaagtac 605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125692 TACCCTATCACTGGCCTCAGGCCCCCGCAGGCTGCCTCTTCCCAACCTCCTTGGAAGTAC 125751

Qy    606 ccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacaca 661
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125752 CCCTGTTTCTTAAACAATTATTTAAGTGTACGTGTATTATTAAACTGATGAACACA 125807


RESULT   2
AF152099
LOCUS       AF152099     1047 bp    mRNA            PRI       27-JAN-2000
DEFINITION  Homo sapiens interleukin 17C mRNA, complete cds.
ACCESSION   AF152099
VERSION     AF152099.1  GI:6759914
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 1047)
  AUTHORS   Li,H., Chen,J., Huang,A., Stinson,J., Heldens,S., Foster,J.,
            Dowd,P., Gurney,A.L. and Wood,W.I.
  TITLE     Cloning and characterization of IL-17B and IL-17C, two new members
            of the IL-17 cytokine family
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (2), 773-778 (2000)
  MEDLINE   20105548
REFERENCE   2  (bases 1 to 1047)
  AUTHORS   Li,H., Chen,J., Stinson,J. and Wood,W.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-1999) Molecular Biology, Genentech, Inc., 1 DNA
            Way, S. San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .1047
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2q31"
                     /clone="DNA62377"
     CDS             50. .643
                     /note="IL-17C"
                     /codon_start=1
                     /product="interleukin 17C"
                     /protein_id="AAF28105.1"
                     /db_xref="GI:6759915"
                     /translation="MTLLPGLLFLTWLHTCLAHHDPSLRGHPHSHGTPHCYSAEELPL
                     GQAPPHLLARGAKWGQALPVALVSSLEAASHRGRHERPSATTQCPVLRPEEVLEADTH
                     QRSISPWRYRVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR
                     RPCSRDGSGLFTPGAFAFHTEFIHVPVGCTCVLPRSV"
BASE COUNT      177 a    381 c    280 g    209 t
ORIGIN
```

```
  Query Match             90.8%;  Score 640.4;  DB 34;  Length 1047;
  Best Local Similarity   98.3%;  Pred. No. 1.8e-112;
  Matches  655;  Conservative    2;  Mismatches    8;  Indels    1;  Gaps    1;

Qy      5 cgagtggacacggatgaggaccgctatccacagaagctggccttcgccgagtgcctgtgc 64
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 CGTGTGGACACGGATGAGGACCGCTATCCACAGAAGCTGGCCTTCGCCGAGTGCCTGTGC 442

Qy     65 agaggctgtatcgatgcacggacgggccgcgagacagctgcgctcaactccgtgcggctg 124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 AGAGGCTGTATCGATGCACGGACGGGCCGCGAGACAGCTGCGCTCAACTCCGTGCGGCTG 502

Qy    125 ctccagagcctgctggtgctgcgccgccggccctgctcccgcgacggctcggggctcccc 184
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    503 CTCCAGAGCCTGCTGGTGCTGCGCCGCCGGCCCTGCTCCCGCGACGGCTCGGGGCTCCCC 562

Qy    185 acacctggggcctttgccttccacaccgagttcatccacgtccccgtcggctgcacctgc 244
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    563 ACACCTGGGGCCTTTGCCTTCCACACCGAGTTCATCCACGTCCCCGTCGGCTGCACCTGC 622

Qy    245 gtgctgccccgttcagtgtgaccgccaaggccgtggggcccttagactggacacgtgtgc 304
          |||||||||||||||||||||| |||||||||||||| ||||||||||||||||||||||
Db    623 GTGCTGCCCCGTTCAGTGTGACCGCCGAGGCCGTGGGGCCCCTAGACTGGACACGTGTGC 682

Qy    305 tccccagagggcacccccctatttatgtgtatttattgttatttatatgcctccccaaca 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 TCCCCAGAGGGCACCCCCTATTTATGTGTATTTATTGTTATTTATATGCCTCCCCCAACA 742

Qy    365 ctacccttggggtctgggcatcccccgtgtctggaggacagccccccactgttctcctca 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 CTACCCTTGGGGTCTGGGCATTCCCCGTGTCTGGAGGACAGCCCCCCACTGTTCTCCTCA 802

Qy    425 tctccagcctcagtagttgggggtwgaaggagctcagcacctcttccagcccttaaagct 484
          |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db    803 TCTCCAGCCTCAGTAGTTGGGGGTAGAAGGAGCTCAGCACCTCTTCCAGCCCTTAAAGCT 862

Qy    485 gcagaaaaggtgtcacacggctgcctgtaccttggytccctgtcctgctcccggcttccc 544
          |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db    863 GCAGAAAAGGTGTCACACGGCTGCCTGTACCTTGGCTCCCTGTCCTGCTCCCGGCTTCCC 922
```

```
Qy    545 ttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttggaagta 604
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    923 TTACCCTATCACTGGCCTCAGG-CCCCGCAGGCTGCCTCTTCCCAACCTCCTTGGAAGTA 981

Qy    605 cccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaa 664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 CCCCTGTTTCTTAAACAATTATTTAAGTGTACGTGTATTATTAAACTGATGAACACATCC 1041

Qy    665 aaaaaa 670
          ||||
Db   1042 CCAAAA 1047
```

```
RESULT   3
AF082526
LOCUS       AF082526     1315 bp    mRNA            ROD       07-SEP-1998
DEFINITION  Mus musculus MEK binding partner 1 (Mp1) mRNA, complete cds.
ACCESSION   AF082526
VERSION     AF082526.1  GI:3549604
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1315)
  AUTHORS   Schaeffer,H.J., Catling,A.D., Eblen,S.T., Collier,L.S., Krauss,A.
            and Weber,M.J.
  TITLE     MP1: A MEK binding partner that enhances enzymatic activation of
            the MAP Kinase cascade
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1315)
  AUTHORS   Schaeffer,H.J., Catling,A.D. and Weber,M.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-AUG-1998) Microbiology, University of Virginia, JPA
            1300, Charlottesville, VA 22908, USA
FEATURES             Location/Qualifiers
     source          1. .1315
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene            1. .1315
                     /gene="Mp1"
     CDS             147. .521
                     /gene="Mp1"
                     /codon_start=1
                     /product="MEK binding partner 1"
                     /protein_id="AAC34591.1"
                     /db_xref="GI:3549605"
                     /translation="MADDLKRFLYKKLPSVEGLHAIVVSDRDGVPVIKVANDSAPEHA
                     LRPGFLSTFALATDQGSKLGLSKNKSIICYYNTYQVVQFNRLPLVVSFIASSSANTGL
                     IVSLEKELAPLFEELIKVVEVS"
BASE COUNT      395 a    273 c    300 g    347 t
ORIGIN
```

```
  Query Match              9.1%;  Score 64.2;  DB 74;  Length 1315;
  Best Local Similarity   74.3%;  Pred. No. 0.0094;
  Matches   81;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy    597 tggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatga 656
          ||  || |   |  ||| ||| |||  | | ||| |  |   ||||| |||||   |  |
Db   1207 TGACAGGATGGCCATTTTTTATACATGTGTATAAATAAAATAGTATTTTTAAAGGTAAAA 1266

Qy    657 acacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
          | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1267 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1315
```

```
RESULT   4
CPCRM7
LOCUS       CPCRM7       989 bp    mRNA            PLN       09-JUL-1997
DEFINITION  C.pteroides mRNA for CRM7 protein.
ACCESSION   Y08243
VERSION     Y08243.1  GI:2252479
```

```
KEYWORDS    CRM7 gene; MADS domain; transcription factor.
SOURCE      Ceratopteris pteridoides.
  ORGANISM  Ceratopteris pteridoides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Filicophyta; Filicopsida; Filicales; Pteridaceae;
            Ceratopteris.
REFERENCE   1  (bases 1 to 989)
  AUTHORS   Theissen,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-1996) G. Theissen, MPI fuer Zuechtungsforschung,
            Molekulare Pflanzengenetik, Carl-Von-Linne-Weg 10, 50829 Koeln,
            D-50829, FRG
REFERENCE   2  (bases 1 to 989)
  AUTHORS   Munster,T., Pahnke,J., Di Rosa,A., Kim,J.T., Martin,W., Saedler,H.
            and Theissen,G.
  TITLE     Floral homeotic genes were recruited from homologous MADS-box genes
            preexisting in the common ancestor of ferns and seed plants
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 94 (6), 2415-2420 (1997)
  MEDLINE   97225967
FEATURES             Location/Qualifiers
     source          1. .989
                     /organism="Ceratopteris pteridoides"
                     /db_xref="taxon:58167"
     gene            1. .583
                     /gene="CRM7"
     CDS             <1. .583
                     /gene="CRM7"
                     /codon_start=2
                     /product="putative MADS domain transcription factor"
                     /protein_id="CAA69411.1"
                     /db_xref="GI:2252480"
                     /db_xref="SPTREMBL:O23768"
                     /translation="RTGLLKKAYELSVLCDAEVAVIVFSSTGRLSEFATPSMPKMLEK
                     YHRAIQRSQGNEHSLQNIEHWQHEAINLRKRIQLLQLRQSHLMGENLTCFQLKDLDIL
                     ESRLSVTVEKIREEKMRRFHVHAEEMHKQECLLHQENTLLKLKIENVESMKHSAHVID
                     LLSDDDCMLESPKISQESGLKDSRFSLQLGLKP"
BASE COUNT      346 a    170 c    197 g    276 t
ORIGIN
```

```
  Query Match             8.8%;  Score 61.8;  DB 7;  Length 989;
  Best Local Similarity   69.4%;  Pred. No. 0.027;
  Matches   84;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy     585 tcccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtatta 644
           | |  |||| | || | |||| | |  || ||| | | ||  |   |  |  || |   |
Db     866 TTCTCACCTGCCTGTACGTACGCTTTATTATTACAAAGTTGGTGTCGAATTTGTTTTAAA 925

Qy     645 ttaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 704
             |||   |  || | ||||||||||||||||||||||||||||||||||||||||||
Db     926 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 985

Qy     705 a 705
           |
Db     986 A 986
```

```
RESULT   5
AF065135
LOCUS       AF065135     1554 bp    mRNA            VRT       12-MAR-1999
DEFINITION  Xenopus laevis derriere mRNA, complete cds.
ACCESSION   AF065135
VERSION     AF065135.1  GI:4405983
KEYWORDS    .
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 1554)
  AUTHORS   Sun,B.I., Bush,S.M., Collins-Racie,L.A., LaVallie,E.R.,
            DiBlasio-Smith,E.A., Wolfman,N.M., McCoy,J.M. and Sive,H.L.
  TITLE     derriere: a TGF-beta family member required for posterior
            development in Xenopus
  JOURNAL   Development 126 (7), 1467-1482 (1999)
  MEDLINE   99169090
REFERENCE   2  (bases 1 to 1554)
  AUTHORS   Sun,B.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-1998) Whitehead Institute, 9 Cambridge Center,
            Cambridge, MA 02142, USA
FEATURES             Location/Qualifiers
     source          1. .1554
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
     CDS             55. .1119
                     /function="posterior determination"
                     /note="TGFb family member"
                     /codon_start=1
                     /product="derriere"
                     /protein_id="AAD19837.1"
                     /db_xref="GI:4405984"
                     /translation="MAELWLSLSCMFSLLLLTNSSPLTFQERMLLKALGLNTRPNPIA
                     PAPVPKSLRDIFEKGINQDNPCMMEGFGVPGNIVRSYRDQGTIAAIEEPQGSLCLKKF
                     LFFDLSAVENKEQLTLGQLEIKFKHNTYYGQQFHLRLYRTLQLSLKGMRDSKMNRKLL
                     VTQSFRLLHKSLYFNLTKVAEDWKNPEKNMGLILEIYASSELAGGNRSFVVCEPIQSF
                     IYTSLLTVSLDPSNCKTQRAKRSTHSSPPTPSNICKKRRLYIDFKDVGWQNWVIAPRG
                     YMANYCHGECPYPLTEMLRGTNHAVLQTLVHSVEPENTPLPCCAPTKLSPISMLYYDN
                     NDNVVLRHYEDMVVDECGCK"
BASE COUNT      503 a    340 c    309 g    402 t
ORIGIN
```

```
  Query Match             8.7%;  Score 61.4;  DB 4;  Length 1554;
  Best Local Similarity   72.1%;  Pred. No. 0.032;
  Matches   80;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy     595 cttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgat 654
           |||| ||||||  ||| || ||  | |   || | | |  ||    |  |  |||   |
Db    1437 CTTGCAAGTACAGCTGATTGTTGGAAATAAATGTGATTTAACCCAAAAAAAAAAAAAAAA 1496

Qy     655 gaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
            || | ||||||||||||||||||||||||||||||||||||||||||||||
Db    1497 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1547
```

```
RESULT   6
HSM801506
LOCUS       HSM801506    1221 bp    mRNA            PRI       18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp434J1618 (from clone DKFZp434J1618);
            partial cds.
ACCESSION   AL133637
VERSION     AL133637.1  GI:6599285
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1221)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by MediGenomix (Martinsried/Germany) within the cDNA
            sequencing consortium of the German Genome Project. This clone
            (DKFZp434J1618) is available at the RZPD in Berlin. Please contact
            the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .1221
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /clone="DKFZp434J1618"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
                     /tissue_type="testis"
                     /map="22q12.1-qter"
     gene            1. .968
                     /gene="DKFZp434J1618"
     CDS             <1. .968
                     /gene="DKFZp434J1618"
                     /note="KIAA0852"
                     /codon_start=3
                     /product="hypothetical protein"
                     /protein_id="CAB63760.1"
                     /db_xref="GI:6599286"
                     /translation="PKVIKTPVVKKTESPIKLSPATPSRKRSVAVSDEEEVEEEAERR
                     KERCKRGRFVVKEEKKDSNELSDSAGEEDSADLKRAQKDKGLHVEVRVNREWYTGRVT
                     AVEVGKHVVRWKVKFDYVPTDTTPRDRWVEKGSEDVRLMKPPSPEHQSLDTQQEGGEE
                     EVGPVAQQAIAVAEPSTSECLRIEPDTTALSTNHETIDLLVQILRNCLRYFLPPSFPI
                     SKKQLSAMNSDELISFPLKEYFKQYEVGLQNLCNSYQSRADSRAKASEESLRTSERKL
                     RETEEKLQKLRTNIVALLQKVQEDIDINTDDELDAYIEDLITKGD"
     polyA_site      1153
BASE COUNT      365 a    283 c    347 g    226 t
ORIGIN


  Query Match             8.7%;  Score 61.2;  DB 69;  Length 1221;
  Best Local Similarity   73.6%;  Pred. No. 0.035;
  Matches   78;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy     600 aagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaaca 659
           |||   ||||| |||||  |   ||  |||||| ||    ||||  |    | |   | |
Db    1102 AAGATACCCTGATTCTTTGAACCTTCCTTAAGTTTATAAATATTTATTTTTTAAAAGAAA 1161

Qy     660 caaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    1162 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1207


RESULT   7
HSM800438
LOCUS       HSM800438     3469 bp    mRNA            PRI       18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp586M121 (from clone DKFZp586M121);
            partial cds.
ACCESSION   AL050138
VERSION     AL050138.1  GI:4884347
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3469)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by MediGenomix (Martinsried/Germany) within the cDNA
            sequencing consortium of the German Genome Project. This clone
            (DKFZp586M121) is available at the RZPD in Berlin. Please contact
            the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES              Location/Qualifiers
     source           1. .3469
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="DKFZp586M121"
                      /clone_lib="586 (synonym: hute1). Vector pSport1; host
                      DH10B; sites NotI + SalI/MluI"
                      /dev_stage="adult"
                      /tissue_type="uterus"
     gene             1. .2981
                      /gene="DKFZp586M121"
     CDS              <1. .2981
                      /gene="DKFZp586M121"
                      /note="weak similarity to multimerin"
                      /codon_start=3
                      /product="hypothetical protein"
                      /protein_id="CAB43287.1"
                      /db_xref="GI:4884348"
                      /translation="TRRGFSLYTGSSGALSPGGPQAQIAPRPASRHRNWCAYVVTRTV
                      SCVLEDGVETYVKYQPCAWGQPQCPQSIMYRRFLRPRYRVAYKTVTDMEWRCCQGYGG
                      DDCAESPAPALGPASSTPRPLARPARPNLSGSSAGSPLSGLGGEGPGESEKVQQLEEQ
                      VQSLTKELQGLRGVLQGLSGRLAEDVQRAVETAFNGRQQPADAAARPGVHETLNEIQH
                      QLQLLDTRVSTHDQELGHLNNHHGGSSSSGGSRAPAPASAPPGPSEELLRQLEQRLQE
                      SCSVCLAGLDGFRRQQQEDRERLRAMEKLLASVEERQRHLAGLAVGRRPPQECCSPEL
                      GRRLAELERRLDVVAGSVTVLSGRRGTELGGAAGQGGHPPGYTSLASRLSRLEDRFNS
                      TLGPSEEQEESWPGAPGGLSHWLPAARGRLEQLGGLLANVSGELGGRLDLLEEQVAGA
                      MQACGQLCSGAPGEQDSQVSEILSALERRVLDSEGQLRLVGSGLHTVEAAGEARQATL
                      EGLQEVVGRLQDRVDAQDETAAEFTLRLNLTAARLGQLEGLLQAHGDEGCGACGGVQE
                      ELGRLRDGVERCSCPLLPPRGPGAGPGVGGPSRGPLDGFSVFGGSSGSALQALQGELS
                      EVILSFSSLNDSLNELQTTVEGQGADLADLGATKDRIISEINRLQQEATEHATESEER
                      FRGLEEGQAQAGQCPSLEGRLGRLEGVCERLDTVAGGLQGLREGLSRHVAGLWAGLRE
                      TNTTSQMQAALLEKLVGGQAGLGRRLGALNSSLQLLEDRLHQLSLKDLTGPAGEAGPP
                      GPPGLQGPPGPAGPPGSPGKDGQEGPIGPPGPQGEQGVEGAPAAPVPQVAFSAALSLP
                      RSEPGTVPFDRVLLNDGGYYDPETGVFTAPLAGRYLLSAVLTGHRHEKVEAVLSRSNQ
                      GVARVDSGGYEPEGLENKPVAESQPSPGTLGVFSLILPLQAGDTVCVDLVMGQLAHSE
                      EPLTIFSGALLYGDPELEHA"
     polyA_signal     3340. .3345
     polyA_site       3361
BASE COUNT      672 a   1058 c   1220 g    519 t
ORIGIN


  Query Match             8.6%;  Score 60.8;  DB 69;  Length 3469;
  Best Local Similarity   57.5%;  Pred. No. 0.042;
  Matches  107;  Conservative    1;  Mismatches   78;  Indels    0;  Gaps    0;

Qy     520 ytccctgtcctgctcccggcttcccttaccctatcactggcctcaggcccccgcaggctg 579
           :|||    ||| | |  | | |||    |  |  |   |||  ||||| ||| |  || |
Db    3248 CTCCGCTCCCTCCACTGGCCCTCCAGGTCGATTCCCTGGGCTCCAGGCTCCCCCGCGCGG 3307

Qy     580 cctcttcccaacctccttggaagtaccccgtttcttaaacaattatttaagtgtacgtg 639
           | |  |||| |  | |   ||      |  |      | ||| ||  ||       |
Db    3308 GCGCCGCCCACCGCCATACTAAACGATCGAGGAATAAAGACACTTGGTTTTTCCAAAAAA 3367

Qy     640 tattattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 699
           |  |  |||   |  || | ||||||||||||||||||||||||||||||||||||||
Db    3368 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3427

Qy     700 aaaaaa 705
           ||||||
Db    3428 AAAAAA 3433


RESULT   8
AF230496
LOCUS       AF230496     6410 bp    mRNA            PRI       24-JUL-2000
DEFINITION  Homo sapiens NADPH thyroid oxidase 2 (THOX2) mRNA, complete cds.
ACCESSION   AF230496
VERSION     AF230496.1  GI:8163927
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6410)
  AUTHORS   De Deken,X., Wang,D., Many,M.C., Costagliola,S., Libert,F.,
            Vassart,G., Dumont,J.E. and Miot,F.
  TITLE     Cloning of two human thyroid cDNAs encoding new members of the
            NADPH oxidase family
```

```
  JOURNAL   J. Biol. Chem. 275 (30), 23227-23233 (2000)
  MEDLINE   20390085
REFERENCE   2  (bases 1 to 6410)
  AUTHORS   De Deken,X., Wang,D., Many,M.-C., Costagliola,S., Libert,F.,
            Dumont,J.E., Vassart,G. and Miot,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-2000) I.R.I.B.H.N, U.L.B, Route de Lennik, 808,
            Bruxelles, Brabant 1070, Belgium
FEATURES             Location/Qualifiers
     source          1. .6410
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q15.3"
     gene            1. .6410
                     /gene="THOX2"
     CDS             186. .4832
                     /gene="THOX2"
                     /note="flavoprotein NADPH oxidase; thyroid H2O2 generating
                     system"
                     /codon_start=1
                     /product="NADPH thyroid oxidase 2"
                     /protein_id="AAF73922.1"
                     /db_xref="GI:8163928"
                     /translation="MLRARPEALMLLGALLTGSLGPSGSQDALSLPWEVQRYDGWFNN
                     LRHHERGAVGCRLQRRVPANYADGVYQALEEPQLPNPRRLSNAATRGIAGLPSLHNRT
                     VLGVFFGYHVLSDVVSVETPGCPAEFLNIRIPPGDPVFDPDQRGDVVLPFQRSRWDPE
                     TGRSPSNPRDLANQVTGWLDGSAIYGSSHSWSDALRSFSGGQLASGPDPAFPRDSQNP
                     LLMWAAPDPATGQNGPRGLYAFGAERGNREPFLQALGLLWFRYHNLWAQRLARQHPDW
                     EDEELFQHARKRVIATYQNIAVYEWLPSFLQKTLPEYTGYRPFLDPSISPEFVVASEQ
                     FFSTMVPPGVYMRNASCHFRKVLNKGFQSSQALRVCNNYWIRENPNLNSTQEVNELLL
                     GMASQISELEDNIVVEDLRDYWPGPGKFSRTDYVASSIQRGRDMGLPSYSQALLAFGL
                     DIPRNWSDLNPNVDPQVLEATAALYNQDLSQLELLLGGLLESHGDPGPLFSAIVLDQF
                     VRLRDGDRYWFENTRNGLFSKKEIEDIRNTTLRDVLVAVINIDPSALQPNVFVWHKGA
                     PCPQPKQLTTDGLPQCAPLTVLDFFEGSSPGFAITIIALCCLPLVSLLLSGVVAYFRG
                     REHKKLQKKLKESVKKEAAKDGVPAMEWPGPKERSSPIIIQLLSDRCLQVLNRHLTVL
                     RVVQLQPLQQVNLILSNNRGCRTLLLKIPKEYDLVLLFSSEEERGAFVQQLWDFCVRW
                     ALGLHVAEMSEKELFRKAVTKQQRERILEIFFRHLFAQVLDINQADAGTLPLDSSQKV
                     REALTCELSRAEFAESLGLKPQDMFVESMFSLADKDGNGYLSFREFLDILVVFMKGSP
                     EDKSRLMFTMYDLDENGFLSKDEFFTMMRSFIEISNNCLSKAQLAEVVESMFRESGFQ
                     DKEELTWEDFHFMLRDHDSELRFTQLCVKGGGGGGNGIRDIFKQNISCRVSFITRTPG
                     ERSHPQGLGPPAPEAPELGGPGLKKRFGKKAAVPTPRLYTEALQEKMQRGFLAQKLQQ
                     YKRFVENYRRHIVCVAIFSAICVGVFADRAYYYGFALPPSDIAQTTLVGIILSRGTAA
                     SVSFMFSYILLTMCRNLITFLRETFLNRYVPFDAAVDFHRWIAMAAVVLAILHSAGHA
                     VNVYIFSVSPLSLLACIFPNVFVNDGSKLPQKFYWWFFQTVPGMTGVLLLLVLAIMYV
                     FASHHFRRRSFRGFWLTHHLYILLYALLIIHGSYALIQLPTFHIYFLVPAIIYGGDKL
                     VSLSRKKVEISVVKAELLPSGVTYLQFQRPQGFEYKSGQWVRIACLALGTTEYHPFTL
                     TSAPHEDTLSLHIRAVGPWTTRLREIYSSPKGNGCAGYPKLYLDGPFGEGHQEWHKFE
                     VSVLVGGGIGVTPFASILKDLVFKSSLGSQMLCKKIYFIWVTRTQRQFEWLADIIQEV
                     EENDHQDLVSVHIYVTQLAEKFDLRTTMLYICERHFQKVLNRSLFTGLRSITHFGRPP
                     FEPFFNSLQEVHPQVRKIGVFSCGPPGMTKNVEKACQLVNRQDRAHFMHHYENF"
     misc_feature    483. .485
                     /gene="THOX2"
                     /note="putative glycosylation site"
     misc_feature    1227. .1229
                     /gene="THOX2"
                     /note="putative glycosylation site"
     misc_feature    1329. .1331
                     /gene="THOX2"
                     /note="putative glycosylation site"
     misc_feature    1548. .1550
                     /gene="THOX2"
                     /note="putative glycosylation site"
     misc_feature    1794. .1796
                     /gene="THOX2"
                     /note="putative glycosylation site"
     misc_feature    2679. .2717
                     /gene="THOX2"
                     /note="Region: EF-hand motif"
     misc_feature    2787. .2825
                     /gene="THOX2"
                     /note="Region: EF-hand motif"
     misc_feature    3435. .3437
                     /gene="THOX2"
                     /note="Arginine: heme binding; binding site"
     misc_feature    3564. .3566
                     /gene="THOX2"
                     /note="Histidine: heme binding; binding site"
     misc_feature    3606. .3608
                     /gene="THOX2"
                     /note="Histidine: heme binding; binding site"
     misc_feature    3849. .3851
                     /gene="THOX2"
                     /note="Histidine: heme binding; binding site"
     misc_feature    3888. .3890
                     /gene="THOX2"
                     /note="Histidine: heme binding; binding site"
     misc_feature    4131. .4151
                     /gene="THOX2"
                     /note="Region: FAD binding motif"
     misc_feature    4341. .4361
                     /gene="THOX2"
                     /note="Region: NADPH binding motif"
     misc_feature    4434. .4448
                     /gene="THOX2"
                     /note="Region: NADPH binding motif"
     misc_feature    4632. .4661
                     /gene="THOX2"
                     /note="Region: NADPH binding motif"
     misc_feature    4725. .4742
                     /gene="THOX2"
                     /note="Region: NADPH binding motif"
BASE COUNT     1441 a   1871 c   1686 g   1412 t
ORIGIN


  Query Match           8.6%;  Score 60.8;  DB 34;  Length 6410;
  Best Local Similarity   80.7%;  Pred. No. 0.042;
  Matches   71;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy     618 aacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaaaaaaaaa 677
           || |||| | | ||||| | |||| | ||| | || | ||||||||||||||||||
Db    6305 AAAAATTCTGTGAGTGTTCTTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6364

Qy     678 aaaaaaaaaaaaaaaaaaaaaaaaaa 705
           ||||||||||||||||||||||||||
Db    6365 AAAAAAAAAAAAAAAAAAAAAAAAAA 6392


RESULT   9
AK025967
LOCUS       AK025967     1478 bp    mRNA            PRI       29-SEP-2000
DEFINITION  Homo sapiens cDNA: FLJ22314 fis, clone HRC05250.
ACCESSION   AK025967
VERSION     AK025967.1  GI:10438649
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens primary human renal epithelial cells cDNA to mRNA,
            clone_lib:HRC clone:HRC05250.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished (2000)
REFERENCE   2  (bases 1 to 1478)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio
            Sugano, Institute of Medical Science, University of Tokyo,
            Laboratory of Genome Structure Analysis, Human Genome Center;
            Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
```

```
          Fax:81-3-5449-5416)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
          International Trade and Industry of Japan; cDNA full insert
          sequencing: Research Association for Biotechnology; cDNA library
          construction, 5'- & 3'-end one pass sequencing: Departent of
          Virology and Human Genome Center, Institute of Medical Science,
          University of Tokyo (partly supported by Science and Technology
          Agency).
FEATURES             Location/Qualifiers
     source          1. .1478
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="primary human renal epithelial cells"
                     /clone="HRC05250"
                     /clone_lib="HRC"
                     /note="cloning vector pME18SFL3"
BASE COUNT      501 a    271 c    343 g    363 t
ORIGIN


  Query Match             8.6%;  Score 60.4;  DB 34;  Length 1478;
  Best Local Similarity   69.5%;  Pred. No. 0.05;
  Matches   82;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy     588 caacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattatta 647
           ||  |||  ||     |    ||  ||||||   || ||  |||  ||| |  |  |||
Db    1327 CAGACTCTGTGTGTACAGTTATGGGTCTTAATTTATAATAAAAGGTTACATTGAAAATTG 1386

Qy     648 aactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           ||   |  || | ||||||||||||||||||||||||||||||||||||||||||||
Db    1387 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1444


RESULT  10
HSM802263
LOCUS       HSM802263    1678 bp    mRNA            PRI       18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp434C2019 (from clone DKFZp434C2019);
            partial cds.
ACCESSION   AL137530
VERSION     AL137530.1  GI:6808201
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1678)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by MediGenomix (Martinsried/Germany) within the cDNA
            sequencing consortium of the German Genome Project. This clone
            (DKFZp434C2019) is available at the RZPD in Berlin. Please contact
            the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .1678
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434C2019"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
                     /tissue_type="testis"
     gene            1. .436
                     /gene="DKFZp434C2019"
     CDS             <1. .436
                     /gene="DKFZp434C2019"
                     /note="unknown"
                     /codon_start=2
                     /product="hypothetical protein"
                     /protein_id="CAB70794.1"
                     /db_xref="GI:6808202"
                     /translation="ALQGPKTQHSIHTVTCKSPRQKEDRSPKPPQAPQHPEEHGRQSH
                     SSSSFASGTLQDMWRLLDLGSSPSGVTSQGDSTPGRRLPSPGAAHSWQLARATLSEGH
                     VESQHPACLLHPEIPALWLSNAWSLFTSHLAPECFPGYGAKP"
     polyA_site      1618
BASE COUNT      380 a    452 c    432 g    414 t
ORIGIN


  Query Match             8.6%;  Score 60.4;  DB 69;  Length 1678;
  Best Local Similarity   71.8%;  Pred. No. 0.05;
  Matches   79;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy     596 ttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatg 655
           ||  | || |   | |||  |    |  || || | ||| | ||| || | |||   |
Db    1569 TTTTATGTTCATATTTTTTATTTTTATGTAATTTAATGTCCTTGTTTTTTAAAAAAAAAA 1628

Qy     656 aacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           || | ||||||||||||||||||||||||||||||||||||||||||||||
Db    1629 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1678


RESULT  11
AF254119
LOCUS       AF254119     2358 bp    mRNA            MAM       27-APR-2000
DEFINITION  Ovis aries angiotensin II type 1 receptor mRNA, complete cds.
ACCESSION   AF254119
VERSION     AF254119.1  GI:7650484
KEYWORDS    .
SOURCE      sheep.
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 2358)
  AUTHORS   Bird,I.M. and Millican,D.S.
  TITLE     Isolation of a Full Length Ovine Angiotensin II Type-1 Receptor
            (AT1-R) cDNA
  JOURNAL   Endocr. Res. (2000) In press
REFERENCE   2  (bases 1 to 2358)
  AUTHORS   Bird,I.M. and Millican,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-APR-2000) 7E Meriter, UW Madison, 202 South Park
            Street, Madison, WI 53715, USA
FEATURES             Location/Qualifiers
     source          1. .2358
                     /organism="Ovis aries"
                     /db_xref="taxon:9940"
                     /tissue_type="adrenal cortex"
     CDS             322. .1401
                     /note="AT1-R"
                     /codon_start=1
                     /product="angiotensin II type 1 receptor"
                     /protein_id="AAF66063.1"
                     /db_xref="GI:7650485"
                     /translation="MILNSSTEDGIKRIQDDCPKAGRHNYIFIMIPTLYSIIFVVGLF
                     GNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLTLPLWAVYTAMEYRWPFGNYLCK
                     IASGSVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTCIIIWLLAGLAS
                     LPTIIHRNVFFIENTNITVCAFHYESQNSTLPVGLGLTKNILGFLFPFLIILTSYTLI
                     WKTLKKAYEIQKNKPRKDDIFKIILAIVLFFFFSWVPHQIFTFMDVLIQLGLIRDCKI
                     EDIVDTAMPITICLAYFNNCLNPLFYGFLGKKFKKYFLQLLKYIPPKAKSHSNLSTKM
                     STLSYRPSENGNSSTKKPAPCTEVE"
BASE COUNT      710 a    511 c    447 g    690 t
ORIGIN


  Query Match             8.5%;  Score 60.2;  DB 3;  Length 2358;
  Best Local Similarity   76.3%;  Pred. No. 0.055;
  Matches   74;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;
```

```
Qy    609 tgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaa 668
          | |||  |   || |  || || |  | || | || |||||| |   || | ||||||||
Db   2253 TATTTAATCTTCATTACTTAAAATAAATGTTTTTTTTTAAACAGGAAAAAAAAAAAAAAA 2312

Qy    669 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
          |||||||||||||||||||||||||||||||||||||
Db   2313 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2349


RESULT  12
CCCHSMR
LOCUS       CCCHSMR      1419 bp    mRNA            PLN       15-APR-1998
DEFINITION  C.chinensis mRNA for chalcone synthase.
ACCESSION   Z67988
VERSION     Z67988.1  GI:3059082
KEYWORDS    chalcone synthase.
SOURCE      Callistephus chinensis.
  ORGANISM  Callistephus chinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
            Asteridae; Asterales; Asteraceae; Callistephus.
REFERENCE   1  (bases 1 to 1419)
  AUTHORS   Henkel,J., Wassenegger,M., Sommer,H. and Forkmann,G.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1419)
  AUTHORS   Henkel,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1995) Henkel J., Technische Universitaet
            Muenchen, Lehrstuhl fuer Zierpflanzenbau, Blumenstrasse 16, 85350
            Freising, Germany
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 1419)
  AUTHORS   Henkel,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1997) Henkel J., Technische Universitaet
            Muenchen, Lehrstuhl fuer Zierpflanzenbau, Blumenstrasse 16, 85350
            Freising, Germany
COMMENT     On Apr 18, 1998 this sequence version replaced gi:1632780.
FEATURES             Location/Qualifiers
     source          1. .1419
                     /organism="Callistephus chinensis"
                     /strain="L 01 (Institute collection)"
                     /db_xref="taxon:13379"
                     /dev_stage="buds and flowers"
                     /tissue_type="petals"
     CDS             33. .1229
                     /EC_number="2.3.1.74"
                     /codon_start=1
                     /product="chalcone synthase"
                     /protein_id="CAA91930.1"
                     /db_xref="GI:3059083"
                     /db_xref="SWISS-PROT:P48385"
                     /translation="MASTIDIAAIREAQRRQGPATILAIGTATPSNCVYQADYPDYYF
                     RITKSEHMVDLKEKFKRMCDKSMIRKRYMHLTEEYLKENPSLCEYMAPSLDARQDVVV
                     VEVPKLGKEAATKAIKEWGQPKSKITHLIFCTTSGVDMPGADYQLTKLLGLRPSVKRF
                     MMYQQGCFAGGTVLRLAKDLAENNKGARVLVVCSEITAVTFRGPNDTHLDSLVGQALF
                     GDGAAAVIVGADPDLTTERPLFEMISAAQTILPDSEGAIDGHLREVGLTFHLLKDVPG
                     LISKNIEKALTQAFSPLGITDWNSIFWIAHPGGPAILDQVELKLGLKEEKMRATRHVL
                     SEYGNMSSACVLFIIDEMRKKSAEDGAATTGEGLDWGVLFGFGPGLTVETVVLHSLPT
                     TMAIAT"
     polyA_signal    1298. .1303
     polyA_signal    1332. .1337
     polyA_site      1357
BASE COUNT      409 a    320 c    333 g    357 t
ORIGIN


  Query Match             8.5%;  Score 60;  DB 7;  Length 1419;
  Best Local Similarity   75.0%;  Pred. No. 0.059;
  Matches   75;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy    606 ccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaa 665
          || | ||| || |  |  ||| |   |  | || || ||| |||   |  || | |||||
Db   1318 CCGTCTTTTTTTATAATATATATTGTTTAATGTTTAATATAAAAAAAAAAAAAAAAAAAA 1377

Qy    666 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
          |||||||||||||||||||||||||||||||||||||||
Db   1378 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1417


RESULT  13
HSM801246
LOCUS       HSM801246    3236 bp    mRNA            PRI       18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp434B2115 (from clone DKFZp434B2115).
ACCESSION   AL122093
VERSION     AL122093.1  GI:6102904
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3236)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp434B2115) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .3236
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434B2115"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
                     /tissue_type="testis"
     polyA_signal    3103. .3108
     polyA_site      3123
BASE COUNT      955 a    758 c    759 g    764 t
ORIGIN


  Query Match             8.5%;  Score 60;  DB 69;  Length 3236;
  Best Local Similarity   72.2%;  Pred. No. 0.06;
  Matches   78;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy    598 ggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaa 657
          | || ||| | |||    | | | | ||| ||    ||| |  | ||| ||| |   |
Db   3064 GTAACTACACTTGTGATATCACCTAGTATATATTAATACATTAAATATGAAAATTCAGTG 3123

Qy    658 cacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
          | ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3124 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3171


RESULT  14
ATPUR2
LOCUS       ATPUR2       1809 bp    mRNA            PLN       03-MAY-1999
DEFINITION  Arabidopsis thaliana mRNA for glycinamide ribonucleotide synthetase
            encoded by PUR2 gene.
ACCESSION   X74766
VERSION     X74766.2  GI:4741182
KEYWORDS    glycinamide ribonucleotide synthetase; phosphoribosylamineglycine
            ligase; pur2 gene.
SOURCE      thale cress.
```

```
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
            eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
            Arabidopsis.
REFERENCE   1  (bases 1 to 1809)
  AUTHORS   Schnorr,K.M., Nygaard,P. and Laloue,M.
  TITLE     Molecular characterization of Arabidopsis thaliana cDNAs encoding
            three purine biosynthetic enzymes
  JOURNAL   Plant J. 6 (1), 113-121 (1994)
  MEDLINE   95004650
REFERENCE   2  (bases 1 to 1809)
  AUTHORS   Schnorr,K.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-AUG-1993) K.M. Schnorr, Inst. National de la
            Recherche, Agronomique-Versailles, Lab. de Biologie Cellulaire,
            Route de Saint Cyr, Versailles F-78026, FRANCE
  REMARK    Revised by [4]
REFERENCE   3  (bases 1 to 1809)
  AUTHORS   Schnorr,K.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-1999) K.M. Schnorr, Inst. National de la
            Recherche, Agronomique-Versailles, Lab. de Biologie Cellulaire,
            Route de Saint Cyr, Versailles F-78026, FRANCE
COMMENT     On May 4, 1999 this sequence version replaced gi:1019914.
FEATURES             Location/Qualifiers
     source          1. .1809
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /tissue_type="whole plant"
                     /clone_lib="lambda YES"
     gene            648. .1750
                     /gene="PUR2"
     CDS             648. .1583
                     /gene="PUR2"
                     /codon_start=1
                     /product="glycinamide ribonucleotide synthetase"
                     /protein_id="CAA52778.2"
                     /db_xref="GI:4741183"
                     /translation="MKNLCHKYNIPTAKYKTFSDASAAKEYIQEQGAPIVIKADGLAA
                     GKGVTVAMELEEAFEAVDSMLVKGVFGSAGCQVVVEEFLEGEEASFFALVDGENAIPL
                     ESAQDHKRVGDGDTGPNTGGMGAYSPAPVLTKELQDFVMESIIHPTVKGMAEEGCKFV
                     GVLFAGLMIEKKSGFLSFIEFNVRFGDPECQVLMMRLESDLAKVLLAACKGELSGVSL
                     DWSKDSAMVVVMASNGYPGSYEKGSIIKNLEEAERVAPGVKVFHAGTGLDSEGNVVAT
                     GGRVLGVTAKGKDLEEARERAYSAVQQINWPGRFL"
     polyA_signal    1732. .1738
                     /gene="PUR2"
     polyA_site      1750
                     /gene="PUR2"
BASE COUNT      503 a    370 c    430 g    506 t
ORIGIN
```

```
  Query Match             8.5%;  Score 59.8;  DB 7;  Length 1809;
  Best Local Similarity   68.9%;  Pred. No. 0.065;
  Matches   82;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy     587 ccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattatt 646
           ||||   | |   |   || ||  | | ||||   || ||| ||    ||   | || ||
Db    1682 CCAAAAACTTAAAACTCACTCCAATCTGTTAAGAGATGATTCAAAATAACCATTTTTGTT 1741

Qy     647 aaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
             ||   |  || | ||||||||||||||||||||||||||||||||||||||||||
Db    1742 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1800


RESULT  15
HSM801960
LOCUS       HSM801960    4173 bp    mRNA            PRI       18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp434L1016 (from clone DKFZp434L1016).
ACCESSION   AL137292
VERSION     AL137292.1  GI:6807752
KEYWORDS    .
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4173)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp434L1016) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .4173
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434L1016"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
                     /tissue_type="testis"
     polyA_signal    4079. .4084
     polyA_site      4096
BASE COUNT      982 a    1221 c    1008 g    962 t
ORIGIN
```

```
  Query Match             8.5%;  Score 59.8;  DB 69;  Length 4173;
  Best Local Similarity   76.8%;  Pred. No. 0.066;
  Matches   73;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy     611 tttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaa 670
           | |||  |||| | | | |||||   | || | |  |||   |  || | ||||||||||
Db    4065 TGTCTATAACAGTGAATAAAGTGCTTGGGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAA 4124

Qy     671 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||
Db    4125 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4159


Search completed: February 24, 2001, 20:41:30
Job time: 8598 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 19:08:07 ; Search time 1413.06 Seconds
(without alignments)
3496.156 Million cell updates/sec

Title: US-09-320-713-1
Perfect score: 705
Sequence: 1 ggcacgagtggacacggatg.....aaaaaaaaaaaaaaaaaaa 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*
44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estpl1:*
70: em_estpl2:*
71: em_estpl3:*
72: em_estpl4:*
73: em_estpl5:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estp16:*
118: em_estp17:*
119: em_estp18:*
120: em_estro14:*
121: em_estro15:*
122: em_estro16:*
123: em_estro17:*
124: em_estro18:*
125: em_estro19:*
126: gb_est58:*
127: gb_est59:*
128: gb_est60:*
129: gb_est61:*
130: gb_est62:*
131: gb_est63:*
132: gb_est64:*
133: gb_est65:*
134: gb_est66:*
135: gb_est75:*
136: gb_est76:*
137: gb_est77:*
138: gb_est78:*
139: gb_est79:*
140: gb_est80:*
141: gb_est81:*
142: gb_est82:*
143: gb_est83:*
144: gb_est84:*
145: gb_est85:*
146: gb_est86:*
147: gb_est87:*
148: gb_gss1:*
149: gb_gss2:*
150: gb_gss3:*
151: gb_gss4:*
152: em_gss1:*
153: em_gss2:*
154: em_gss3:*
155: em_gss4:*
156: gb_gss5:*
157: gb_gss6:*
158: gb_gss7:*
159: gb_gss8:*
160: gb_gss9:*
161: em_gss5:*
162: em_gss6:*
163: em_gss7:*
164: em_gss8:*
165: em_gss9:*
166: em_gss10:*
167: em_gss11:*
168: gb_gss10:*
169: gb_gss11:*
170: em_gss12:*
171: gb_gss12:*
172: gb_gss13:*
173: gb_gss14:*
174: gb_gss15:*
175: gb_gss16:*
176: gb_gss17:*
177: gb_gss18:*
178: gb_gss19:*
179: em_gss13:*
180: gb_gss20:*
181: gb_gss21:*
182: gb_gss22:*
183: gb_gss23:*
184: gb_gss24:*
185: em_gss14:*
186: em_gss15:*
187: em_gss16:*
188: em_gss17:*
189: em_gss18:*
190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	260	36.9	264	109	BE551188	BE551188 7b55g03.x
	2	71	10.1	242	29	AU037868	AU037868 AU037868
c	3	69.8	9.9	369	39	AW090736	AW090736 xc95g02.x
	4	69	9.8	366	111	BE723174	BE723174 192592 MA
c	5	68.8	9.8	359	23	AI673363	AI673363 tw35b08.x
	6	67.8	9.6	199	3	AA198893	AA198893 mu05h09.r
c	7	66.8	9.5	294	22	AI581387	AI581387 to71c06.x
c	8	66.8	9.5	1058	137	BE905366	BE905366 601499028
	9	66	9.4	309	140	C90935	C90935 C90935 Dict
c	10	65.4	9.3	363	92	AW541819	AW541819 C0149H07-
	11	65.2	9.2	394	136	BE808909	BE808909 214216 MA
c	12	64.8	9.2	225	27	AI961414	AI961414 wt22c07.x
c	13	64.8	9.2	550	26	AI925744	AI925744 wo34h11.x
	14	64.6	9.2	400	90	AW394524	AW394524 sh32b03.y
c	15	64.6	9.2	1201	190	CNS016DQ	AL106616 Drosophil
	16	64.4	9.1	245	4	AA274319	AA274319 TgESTzz25
	17	64.4	9.1	308	140	C90985	C90985 C90985 Dict
c	18	64	9.1	263	25	AI818980	AI818980 wj89h11.x
c	19	63.8	9.0	367	87	AW190194	AW190194 xl12g08.x
	20	63.6	9.0	206	29	AU038875	AU038875 AU038875
c	21	63.6	9.0	261	25	AI819976	AI819976 wj88h11.x
	22	63.6	9.0	285	94	AW733216	AW733216 sk71e02.y
c	23	63.6	9.0	322	27	AI973152	AI973152 wr52c02.x
c	24	63.6	9.0	349	20	AI475371	AI475371 tl82e01.x
c	25	63.6	9.0	796	190	CNS0118D	AL099943 Drosophil
c	26	63.4	9.0	230	22	AI559531	AI559531 tq50b09.x
	27	63.4	9.0	348	134	BE058451	BE058451 sn16c06.y
c	28	63.4	9.0	795	39	AW024360	AW024360 wv03c04.x
	29	63.4	9.0	822	108	BE420782	BE420782 HWM002.E0
	30	63.2	9.0	121	39	AW059474	AW059474 fe14e11.y
c	31	63.2	9.0	241	22	AI613038	AI613038 ty06h05.x
	32	63.2	9.0	337	111	BE667471	BE667471 151857 MA
	33	63.2	9.0	371	95	AW786724	AW786724 120133 MA
c	34	63	8.9	120	22	AI590043	AI590043 tr75c04.x
c	35	63	8.9	386	18	AI267454	AI267454 aq65a01.x
	36	62.8	8.9	134	29	AU053905	AU053905 AU053905
	37	62.8	8.9	271	40	AW100582	AW100582 sd57c01.y
c	38	62.8	8.9	838	106	BE319360	BE319360 NF016E02R
	39	62.6	8.9	247	3	AA183005	AA183005 mt78c04.r
	40	62.6	8.9	312	4	AA270657	AA270657 va69a07.r
	41	62.4	8.9	205	40	AW100861	AW100861 sd62c12.y
c	42	62.4	8.9	248	18	AI245008	AI245008 qj94b05.x
	43	62.4	8.9	294	105	BE234043	BE234043 140614 MA
c	44	62.2	8.8	224	19	AI395684	AI395684 MA004543.
c	45	62.2	8.8	338	21	AI538637	AI538637 tp70e10.x

ALIGNMENTS

RESULT 1
BE551188/c
LOCUS BE551188 264 bp mRNA EST 10-AUG-2000
DEFINITION 7b55g03.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3232180 3', mRNA sequence.
ACCESSION BE551188
VERSION BE551188.1 GI:9792880
KEYWORDS EST.
SOURCE human.

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 264)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 259.
FEATURES             Location/Qualifiers
     source          1. .264
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3232180"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
BASE COUNT       66 a     50 c     88 g     60 t
ORIGIN


  Query Match             36.9%;  Score 260;  DB 109;  Length 264;
  Best Local Similarity   98.5%;  Pred. No. 6.4e-46;
  Matches  260;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    402 acagccccccactgttctcctcatctccagcctcagtagttgggggtwgaaggagctcag 461
          ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db    264 ACAGCCCCCCACTGTTCTCCTCATCTCCAGCCTCAGTAGTTGGGGGTAGAAGGAGCTCAG 205

Qy    462 cacctcttccagcccttaaagctgcagaaaaggtgtcacacggctgcctgtaccttggyt 521
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||:|
Db    204 CACCTCTTCCAGCCCTTAAAGCTGCAGAAAAGGTGTCACACGGCTGCCTGAACCTTGGCT 145

Qy    522 ccctgtcctgctcccggcttcccttaccctatcactggcctcaggcccccgcaggctgcc 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 CCCTGTCCTGCTCCCGGCTTCCCTTACCCTATCACTGGCCTCAGGCCCCCGCAGGCTGCC 85

Qy    582 tcttcccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgta 641
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     84 TCTTCCCAACCTCCTTGGAAGTACCCCTGTTTCTTAAACAATTATTTAAGTGAACGTGTA 25

Qy    642 ttattaaactgatgaacacaaaaa 665
          |||||||||||||||||||||||||
Db     24 TTATTAAACTGATGAACACAAAAA 1


RESULT   2
AU037868
LOCUS       AU037868      242 bp    mRNA            EST       29-MAR-1999
DEFINITION  AU037868 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSE387, mRNA sequence.
ACCESSION   AU037868
VERSION     AU037868.1  GI:3984621
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 242)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
FEATURES             Location/Qualifiers
     source          1. .242
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSE387"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT       147 a      5 c      1 g     88 t      1 others
ORIGIN


  Query Match             10.1%;  Score 71;  DB 29;  Length 242;
  Best Local Similarity   84.2%;  Pred. No. 1.7e-05;
  Matches   80;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    611 tttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaa 670
          ||| ||||| |||| ||||| |  |  |||||| ||||| | |  || | ||||||||||
Db    121 TTTATTAAAAAATTTTTAATTTAATTTGTATTTTTAAATTTAAAAAAAAAAAAAAAAAA 180

Qy    671 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
          |||||||||||||||||||||||||||||||||||||
Db    181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 215


RESULT   3
AW090736/c
LOCUS       AW090736      369 bp    mRNA            EST       15-OCT-1999
DEFINITION  xc95g02.x1 NCI_CGAP_Brn35 Homo sapiens cDNA clone IMAGE:2592050 3'
            similar to contains TAR1.t2 TAR1 repetitive element ;, mRNA
            sequence.
ACCESSION   AW090736
VERSION     AW090736.1  GI:6048080
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 369)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
```

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 367.

```
FEATURES             Location/Qualifiers
     source          1. .369
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2592050"
                     /clone_lib="NCI_CGAP_Brn35"
                     /tissue_type="tumor, 5 pooled (see description)"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.33 kb. Tumor types include:
                     meningioma, oligodendroglioma, astrocytoma (grade II),
                     medulloblastoma, astrocytoma (grade IV). Life Technologies
                     catalog #: 11544-012"
BASE COUNT       89 a     86 c     93 g    101 t
ORIGIN


  Query Match              9.9%;  Score 69.8;  DB 39;  Length 369;
  Best Local Similarity   63.3%;  Pred. No. 3.2e-05;
  Matches  107;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy     537 ggcttcccttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctcct 596
           ||   || || |||  |    | |||  || |||| |   |  || |        || ||
Db     191 GGGCCCCTTTCCCCCTTTTTGGCCCTGGGGGCCCCCCCCCCCCCCCCCTTGGCCCCCCC 132

Qy     597 tggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatga 656
           |    | |||| |||  | || |||| |||     |   |  || |||||   |  |
Db     131 TTTTTTTTCCCCTTTTTGATTAAAAATTTTTTTCCCCTTTTGGGTTTTTTAAAAAAAAAA 72

Qy     657 acacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           | | ||||||||||||||||||||||||||||||||||||||||||||||
Db      71 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 23


RESULT   4
BE723174
LOCUS        BE723174      366 bp    mRNA            EST       14-SEP-2000
DEFINITION   192592 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION    BE723174
VERSION      BE723174.1  GI:10124470
KEYWORDS     EST.
SOURCE       cow.
  ORGANISM   Bos taurus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
             Bovidae; Bovinae; Bos.
REFERENCE    1  (bases 1 to 366)
  AUTHORS    Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
             Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
             ,W.W. and Keele,J.W.
  TITLE      Design and use of four pooled tissue normalized cDNA libraries for
             EST discovery in cattle
  JOURNAL    Unpublished (2000)
COMMENT      Contact: Smith TPL
             USDA, ARS, US Meat Animal Research Center
             PO Box 166, Clay Center, NE 68933-0166, USA
             Tel: 402 762 4366
             Fax: 402 762 4390
             Email: smith@email.marc.usda.gov
             Single pass sequencing. Bases called and alt_trimmed with phred
             v0.980904.e. Vector identified by cross_match with the -minscore 18
             and -minmatch 12 options.
             PCR PRimers
             FORWARD: AGGAAACAGCTATGACCAT
             BACKWARD: GTTTTCCCAGTCACGACG
             Plate: 104  row: A  column: 2
             Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .366
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 4BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from day 20 and day 40
                     embryos."
BASE COUNT      164 a     40 c     84 g     78 t
ORIGIN


  Query Match              9.8%;  Score 69;  DB 111;  Length 366;
  Best Local Similarity   74.4%;  Pred. No. 4.8e-05;
  Matches   87;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy     589 aacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaa 648
           || ||  ||| ||   || || |||   ||| ||  || ||  |  || |.||  | |||
Db      95 AAACTAGTTGCAAAACCCACTATTTTGGAAATAAAAATGTATTTTGACATATACAAATAA 154

Qy     649 actgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           | |||  || || ||||||||||||||||||||||||||||||||||||||||||
Db     155 AATGAATAAAACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 211


RESULT   5
AI673363/c
LOCUS        AI673363      359 bp    mRNA            EST       15-DEC-1999
DEFINITION   tw35b08.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2261655 3'
             similar to contains element MSR1 repetitive element ;, mRNA
             sequence.
ACCESSION    AI673363
VERSION      AI673363.1  GI:4853094
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 359)
  AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email: Robert_Strausberg@nih.gov
             Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
             Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html
             Insert Length: 1837   Std Error: 0.00
             Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .359
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2261655"
                     /clone_lib="NCI_CGAP_Ut1"
                     /tissue_type="well-differentiated endometrial
                     adenocarcinoma, 7 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Life Technologies catalog #:
                     11538-014"
BASE COUNT       88 a     80 c     82 g    109 t
ORIGIN
```

```
  Query Match             9.8%;  Score 68.8;  DB 23;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 5.2e-05;
  Matches  112;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy    522 ccctgtcctgctcccggcttcccttaccctatcactggcctcaggcccccgcaggctgcc 581
          ||  | ||| | |||||  ||||     |  |   |  || |   ||||| |   | |
Db    203 CCGGGGCCTCCCCCCGGGGTCCCGGGGTCCCTTTTTTTCCCCCCCCCCCCCCCCCCGGG 144

Qy    582 tcttcccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgta 641
            ||     || ||||||    |   ||  ||| ||  | ||  | ||   | |    | ||
Db    143 GGTTTTTGGCCCCCTTGGGTTTTGGCCCTTTTTTTTTAAAAAAAATTTTTTTTTTTTTA 84

Qy    642 ttattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 701
            |  |||   |  || | ||||||||||||||||||||||||||||||||||||||||
Db     83 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 24

Qy    702 aaaa 705
          ||||
Db     23 AAAA 20
```

```
RESULT   6
AA198893
LOCUS       AA198893      199 bp    mRNA            EST       17-FEB-1997
DEFINITION  mu05h09.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:638561
            5', mRNA sequence.
ACCESSION   AA198893
VERSION     AA198893.1  GI:1795069
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 199)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:390553
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 146.
FEATURES             Location/Qualifiers
     source          1. .199
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:638561"
                     /clone_lib="Soares mouse 3NbMS"
                     /sex="male"
                     /tissue_type="Spleen"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through
                     three rounds of normalization, and was constructed by
                     Bento Soares and M.Fatima Bonaldo."
BASE COUNT       90 a     39 c     23 g     47 t
ORIGIN
```

```
  Query Match             9.6%;  Score 67.8;  DB 3;  Length 199;
  Best Local Similarity   60.7%;  Pred. No. 8.3e-05;
  Matches  111;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy    523 cctgtcctgctcccggcttcccttaccctatcactggcctcaggcccccgcaggctgcct 582
          ||  |||||||   | | | |||  ||       || |||    ||   ||   |  |||
Db     17 CCGTTCCTGCTGGTGTCCTGCCTGCCCAGCAACTTGTCCTTCTTCCATTGCTCTCGACCT 76

Qy    583 cttcccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtat 642
          |||  | ||     || |      ||||| | | ||| |||||    |     |
Db     77 CTTGGCCACAAGGATGCAGTGTATATATTTTCTTGATCTATTTTTTAATAAAAATAAAAA 136

Qy    643 tattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 702
          | ||||   |  || | ||||||||||||||||||||||||||||||||||||||||||
Db    137 GAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 196

Qy    703 aaa 705
          |||
Db    197 AAA 199
```

```
RESULT   7
AI581387/c
LOCUS       AI581387      294 bp    mRNA            EST       14-DEC-1999
DEFINITION  to71c06.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2183722 3',
            mRNA sequence.
ACCESSION   AI581387
VERSION     AI581387.1  GI:4565763
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1083   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 249
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2183722"
                     /clone_lib="NCI_CGAP_Gas4"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
BASE COUNT       79 a     47 c     36 g    132 t
```

```
ORIGIN


  Query Match             9.5%;  Score 66.8;  DB 22;  Length 294;
  Best Local Similarity   75.5%;  Pred. No. 0.00014;
  Matches   83;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy     596 ttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatg 655
           | |  |||| ||||  || || || ||  |||||| | ||  |   | || ||| | | |
Db     113 TGGCCAGTAACCCTTATTTTTTAAAAAAAATTTAATTTTAATTAATTAATAAAAATTAGG 54

Qy     656 aacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      53 GTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4


RESULT   8
BE905366/c
LOCUS       BE905366     1058 bp    mRNA            EST       29-SEP-2000
DEFINITION  601499028F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3901086 5',
            mRNA sequence.
ACCESSION   BE905366
VERSION     BE905366.1  GI:10398560
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1058)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM688  row: g  column: 07
            High quality sequence stop: 154.
FEATURES             Location/Qualifiers
     source          1. .1058
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3901086"
                     /clone_lib="NIH_MGC_70"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      377 a    202 c    278 g    201 t
ORIGIN


  Query Match             9.5%;  Score 66.8;  DB 137;  Length 1058;
  Best Local Similarity   61.5%;  Pred. No. 0.00015;
  Matches  107;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy     532 ctcccggcttcccttaccctatcactggcctcaggcccccgcaggctgcctcttcccaac 591
           | |||  |||  |||    |   ||  || |  | |    ||  ||  | |||||  |
Db     204 CCCCCTTCTTTTCTTCTTTTCGTACCACCCACCCCCACAATTAGTGTGTTTTTTCCCCCC 145

Qy     592 ctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaact 651
           | ||         ||||| ||| ||    |||| |||  | |   | | |  |  |||
Db     144 CCCCCCCCCCCCCCCCCTTTTTTTTTTAAATTTTTTGGGGGGGGGGGAAAAAAAAAAAA 85

Qy     652 gatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |  || | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      84 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 31


RESULT   9
C90935
LOCUS       C90935        309 bp    mRNA            EST       01-MAY-1998
DEFINITION  C90935 Dictyostelium discoideum SS (M.Yoshida) Dictyostelium
            discoideum cDNA clone SSJ328, mRNA sequence.
ACCESSION   C90935
VERSION     C90935.1  GI:3097690
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Yoshida,M.
  TITLE     Developmental cDNA in Dictyostelium discoideum (M.Yoshida)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Motonobu Yoshida
            Research Institute of Food Science
            Kinki University
            Nakamachi 3327, Nara 631, Japan
            Email: yoshida@ews06.nara.kindai.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .309
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSJ328"
                     /clone_lib="Dictyostelium discoideum SS (M.Yoshida)"
                     /dev_stage="slug"
BASE COUNT      128 a     55 c     27 g     94 t      5 others
ORIGIN


  Query Match             9.4%;  Score 66;  DB 140;  Length 309;
  Best Local Similarity   63.0%;  Pred. No. 0.00021;
  Matches  102;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy     544 cttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttggaagt 603
           ||  || ||  |    ||||| |   || | || |   || | |  |   | ||| || |
Db     137 CTCCCCTTACGATGCTCCTCAAGTTTCCTCGGGGTTTATCCTTCACATTCCTTTGCAACT 196

Qy     604 acccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaa 663
           |   || || | || || ||||    |    | || | |||   |  || | |||
Db     197 TTTTCAAATTTTTTAGCAGTTTTTTACAAAAAAAAATTTTTTAAAAAAAAAAAAAAAAA 256

Qy     664 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||||||||
Db     257 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 298


RESULT  10
AW541819/c
LOCUS       AW541819      363 bp    mRNA            EST       31-AUG-2000
DEFINITION  C0149H07-3 Mouse E7.5 Extraembryonic Portion cDNA Library Mus
            musculus cDNA clone C0149H07 3', mRNA sequence.
ACCESSION   AW541819
VERSION     AW541819.1  GI:7184236
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 363)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
            ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
            III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  MEDLINE   20381348
```

```
COMMENT     Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224:6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: C0149  row: H  column: 07
            Seq primer: -21M13 Forward
            High quality sequence stop: 363
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .363
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="C0149H07"
                     /clone_lib="Mouse E7.5 Extraembryonic Portion cDNA
                     Library"
                     /sex="unknown"
                     /dev_stage="7.5dpc Embryo"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
                     Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
                     5 EPC. The double-stranded cDNA was  synthesized by
                     Gibco's kit with an Oligo(dT) primer [NotI primer-adapter
                     from GibcoBRL]
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3']
                     from 0.8ug of mRNA . The double-stranded cDNAs were
                     treated with T4  DNA polymerase and purified by
                     ethanol-precipitation. The cDNAs were  ligated to
                     Lone-linker LL-Sal3 (include Sal1 sequence). The cDNAs
                     were  purified by phenol/chloroform and separated from
                     free linkers by  Centricon 100. Then, cDNAs were amplified
                     by long-range high fidelity  PCR using Takara's Ex Taq
                     polymerase. Then, the cDNAs were purified by
                     phenol/chloroform and by Centricon 100. The cDNAs were
                     digested with SalI  and NotI enzymes. Then, the cDNAs were
                     size selected by Gibco's Size  Fractionation Column. The
                     cDNAs were cloned into SalI/NotI site of  pSPORT1 plasmid
                     vector. The DH10B E. coli host was transformed with the
                     ligation mixture by chemical method. The library was
                     constructed by  Xiaohong Wang."
BASE COUNT       65 a     47 c    130 g    121 t
ORIGIN


  Query Match             9.3%;  Score 65.4;  DB 92;  Length 363;
  Best Local Similarity   66.9%;  Pred. No. 0.00028;
  Matches   93;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy     567 cccccgcaggctgcctcttcccaacctccttggaagtacccctgtttcttaaacaattat 626
           |||||  |  |  ||||  ||||  || ||||   |||  ||| ||||     |  | |
Db     170 CCCCCTTAATCCCCCTCCCCCCATTCTTCTTGCCCGTAAGCCTTTTTCCCCCCCCTTAAA 111

Qy     627 ttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaa 686
             || | | | | | || ||  |   |  || | ||||||||||||||||||||||||
Db     110 AAAATTTTCCTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 51

Qy     687 aaaaaaaaaaaaaaaaa 705
           || | ||||||||||||||
Db      50 AATATAAAAAAAAAAAAA 32


RESULT  11
BE808909
LOCUS       BE808909      394 bp    mRNA            EST       20-SEP-2000
DEFINITION  214216 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE808909
VERSION     BE808909.1  GI:10240021
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 394)
  AUTHORS   Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
            Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
            ,W.W. and Keele,J.W.
  TITLE     Design and use of four pooled tissue normalized cDNA libraries for
            EST discovery in cattle
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 70  row: O  column: 1
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .394
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT      128 a    107 c     89 g     70 t
ORIGIN


  Query Match             9.2%;  Score 65.2;  DB 136;  Length 394;
  Best Local Similarity   63.3%;  Pred. No. 0.00031;
  Matches  100;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy     548 ccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttggaagtaccc 607
           | |||   ||| || ||| ||   |||| |  ||  |||||  |  |||    ||  ||
Db     228 CACTAGGCCTGCCCACAGCCCGAGGCAGTCGCCCCTTTCCCCGCAGCCTATCCAGGCACC 287

Qy     608 ctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaa 667
               | | ||  |  |     || |  ||||   |  |||   |  || | |||||||
Db     288 TGCCCCCGTGAAATAAAAGGAGGGTTTGTGTGTGAAAAAAAAAAAAAAAAAAAAAAAAA 347

Qy     668 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||||
Db     348 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 385


RESULT  12
AI961414/c
LOCUS       AI961414      225 bp    mRNA            EST       20-AUG-1999
DEFINITION  wt22c07.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2508204 3'
            similar to contains element MER22 repetitive element ;, mRNA
            sequence.
ACCESSION   AI961414
VERSION     AI961414.1  GI:5754127
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 225)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 224.

FEATURES Location/Qualifiers
source 1. .225
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2508204"
/clone_lib="NCI_CGAP_Ut1"
/tissue_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

BASE COUNT 59 a 46 c 43 g 77 t
ORIGIN

Query Match 9.2%; Score 64.8; DB 27; Length 225;
Best Local Similarity 64.9%; Pred. No. 0.00037;
Matches 96; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 558 ggcctcaggccccgcaggctgcctcttcccaacctccttggaagtacccctgtttctta 617
Db 148 GGCCCCCCTTCCCCCCCCCCCCCCTCCCCTGGTCCCCCGTCCCTGGGGCTTTTTTTTTA 89

Qy 618 aacaattatttaagtgtacgtgtattattaaaactgatgaacacaaaaaaaaaaaaaaaaa 677
Db 88 AAAAGGTTTTTAAAGGTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 29

Qy 678 aaaaaaaaaaaaaaaaaaaaaaaaaaa 705
Db 28 AAAAAAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 13
AI925744/c
LOCUS AI925744 550 bp mRNA EST 08-MAR-2000
DEFINITION wo34h11.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2457285 3' similar to gb:X66363 SERINE/THREONINE-PROTEIN KINASE PCTAIRE-1 (HUMAN);contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION AI925744
VERSION AI925744.1 GI:5661708
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 550)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1420 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 270.

FEATURES Location/Qualifiers
source 1. .550
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2457285"
/clone_lib="NCI_CGAP_Gas4"
/tissue_type="poorly differentiated adenocarcinoma with signet ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

BASE COUNT 128 a 127 c 165 g 127 t 3 others
ORIGIN

Query Match 9.2%; Score 64.8; DB 26; Length 550;
Best Local Similarity 61.0%; Pred. No. 0.00039;
Matches 105; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 534 cccggcttcccttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacct 593
Db 192 CCCGGCCCCTTGTTCCCTTGTGGCTGGCCCAACCCCCCCCCCTTGGGCTGTTCCCCCCCC 133

Qy 594 ccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgtg 653
Db 132 CCTTGGGCTGTTTTTTGGGGGGTTACCAAGCTTTTTTCCCTCCTTTCAAAAAAAAAAAAA 73

Qy 654 tgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
Db 72 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 21

RESULT 14
AW394524
LOCUS AW394524 400 bp mRNA EST 07-FEB-2000
DEFINITION sh32b03.y1 Gm-c1017 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1017-3366 5', mRNA sequence.
ACCESSION AW394524
VERSION AW394524.1 GI:6912936
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE 1 (bases 1 to 400)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com

```
           Seq primer: -40RP from Gibco
           High quality sequence stop: 395.
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1017-3366"
                     /clone_lib="Gm-c1017"
                     /tissue_type="vegetable buds of field grown plants"
                     /lab_host="XL10-Gold"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from vegatable buds of field grown plants. The cDNA
                     library was prepared using the Stratagene pBluescript II
                     XR library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
                     (dT) sequence with a XhoI restriction site. EcoRI adapters
                     were ligated to the blunt-ended cDNA fragments followed by
                     XhoI digestion. The cDNA fragments were directionally
                     cloned into the EcoRI-XhoI restriction site of the
                     pBluescript vector. The ligated cDNA fragments were
                     transformed into XL10-Gold host cells. This library was
                     constructed by Dr. Randy Shoemaker and Dr. John
                     Erpelding."
BASE COUNT      191 a     64 c     43 g    101 t      1 others
ORIGIN

  Query Match             9.2%;  Score 64.6;  DB 90;  Length 400;
  Best Local Similarity   61.7%;  Pred. No. 0.00042;
  Matches  103;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy     539 cttcccttaccctatcactggcctcaggcccccgcaggctgcctcttcccaacctccttg 598
           || ||   | |||| |||   |   ||      |   | | | |    || ||  |  |
Db     172 CTGCCACGATCCTACCACCTACAATAGTATTTTGTCCGTTTCATTGATCCTACGACTATC 231

Qy     599 gaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaac 658
            |||||   || |   |      |||| || | |||  | |  | |||  ||||   ||
Db     232 CAAGTAAATCTTTCAGTAGTTGGATTAATTCAATGTCTGAGACTAATTTTACTGTAAAAA 291

Qy     659 acaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           | ||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 338


RESULT  15
CNS016DQ/c
LOCUS       CNS016DQ     1201 bp    DNA             GSS       26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN15M17 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106616
VERSION     AL106616.1  GI:5622808
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Plasmid Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Drosophila melanogaster"
                     /plasmid="pBeloBAC11"
                     /db_xref="taxon:7227"
                     /clone_lib="DrosBAC"
                     /clone="BACN15M17"
                     /note="end : SP6"
BASE COUNT      254 a    200 c    292 g    310 t    145 others
ORIGIN

  Query Match             9.2%;  Score 64.6;  DB 190;  Length 1201;
  Best Local Similarity   36.8%;  Pred. No. 0.00044;
  Matches  110;  Conservative   60;  Mismatches  129;  Indels    0;  Gaps    0;

Qy     407 cccccactgttctcctcatctccagcctcagtagttgggggtwgaaggagctcagcacct 466
           ::|:| |: :||||::::|| :   ||||   : |:     :          :||  |  :|
Db    1179 YYCYCTCYCYTCTCYYYHTCCYTCYCCTCTCCDTTYTTYTTYCYTYYCTCYTCYCCTTYT 1120

Qy     467 cttccagcccttaaagctgcagaaaaggtgtcacacggctgcctgtaccttggytccctg 526
           :|     ::||     ||          :    | |     ||    ||:   ::| :|
Db    1119 YTCTTTTYYCTCTYTCCTHTCCYCCYHTYTCTTCTCTCTCCCCCCCCCCYCTCCYCAYTC 1060

Qy     527 tcctgctcccggcttcccttaccctatcactggcctcaggcccccgcaggctgcctcttc 586
           ||   | |||  :: ::  :|:|| |: |:       |    :|  ||      : :  :
Db    1059 TCTCCCCCCCCCYYCMMTAWAHCCCAYAAHAAAAAAAAAHTATSCAACACAAAAAMMMMMM 1000

Qy     587 ccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattatt 646
           ::||   || : ::  :|:|::    |::|  |  :: :|||    ::| : : :|:||
Db     999 MMAATGCCCMYMCMMMCMCHCMYMMMMCHYACCCYYYYYYYTTATHCAHHCYYYHYYTHTT 940

Qy     647 aaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           || || || || | ||||||||||||||||||||||||||||:::|||||||||::| |
Db     939 AATCTAATWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAMMWAAAAAAAAARWAGA 881


Search completed: February 24, 2001, 21:04:27
Job time: 6980 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 19:10:02 ; Search time 86.16 Seconds
(without alignments)
1318.686 Million cell updates/sec

Title: US-09-320-713-1
Perfect score: 705
Sequence: 1 ggcacgagtggacacggatg.....aaaaaaaaaaaaaaaaaaa 705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61.2	8.7	882	2	US-08-909-965C-9	Sequence 9, Appli
2	59.6	8.5	7218	1	US-08-232-463-14	Sequence 14, Appl
3	58.2	8.3	1602	1	US-08-530-950-3	Sequence 3, Appli
4	58.2	8.3	1602	3	US-08-888-429A-3	Sequence 3, Appli
5	58.2	8.3	1897	1	US-08-184-632-1	Sequence 1, Appli
6	57.6	8.2	1582	3	US-08-545-196B-10	Sequence 10, Appl
7	57.6	8.2	1582	3	US-08-545-196B-12	Sequence 12, Appl
c 8	57.4	8.1	467	2	US-08-841-349-18	Sequence 18, Appl
9	57	8.1	2989	5	5378464-1	Patent No. 5378464
10	56.6	8.0	1098	3	US-09-248-335-35	Sequence 35, Appl
11	56.4	8.0	748	1	US-08-361-467B-3	Sequence 3, Appli
12	56.4	8.0	748	1	US-08-484-332C-3	Sequence 3, Appli
13	56.4	8.0	5173	1	US-08-242-677-1	Sequence 1, Appli
14	56	7.9	1364	1	US-08-265-087-3	Sequence 3, Appli
15	56	7.9	1364	1	US-08-621-493-3	Sequence 3, Appli
16	56	7.9	1364	2	US-08-965-688-3	Sequence 3, Appli
17	55.8	7.9	1134	3	US-09-248-335-29	Sequence 29, Appl
18	55.6	7.9	1172	1	US-07-945-288-9	Sequence 9, Appli
19	55.6	7.9	1172	1	US-08-462-831-9	Sequence 9, Appli
20	55.6	7.9	1172	1	US-08-461-809-9	Sequence 9, Appli
21	55.6	7.9	1172	1	US-08-461-441-9	Sequence 9, Appli
22	55.6	7.9	1172	4	PCT-US93-08518-9	Sequence 9, Appli
23	55.4	7.9	1066	1	US-08-157-101A-4	Sequence 4, Appli
24	55.4	7.9	1279	3	US-09-248-335-25	Sequence 25, Appl
25	55.2	7.8	568	1	US-08-582-257-20	Sequence 20, Appl
26	55.2	7.8	568	2	US-08-582-298-20	Sequence 20, Appl
27	54.8	7.8	458	1	US-08-524-757-1	Sequence 1, Appli
28	54.8	7.8	536	1	US-08-341-568-1	Sequence 1, Appli
29	54.8	7.8	536	2	US-08-911-020-1	Sequence 1, Appli
30	54.8	7.8	688	5	5498694-3	Patent No. 5498694
31	54.8	7.8	976	2	US-08-504-459-9	Sequence 9, Appli
32	54.8	7.8	1534	1	US-08-300-903A-6	Sequence 6, Appli
33	54.6	7.7	222	3	US-08-481-190-15	Sequence 15, Appl
34	54.6	7.7	222	4	PCT-US93-00869-15	Sequence 15, Appl
35	54.4	7.7	140	1	US-08-628-417-5	Sequence 5, Appli
36	54.4	7.7	240	1	US-08-628-417-6	Sequence 6, Appli
37	54.4	7.7	1493	1	US-08-340-820-24	Sequence 24, Appl
38	54.4	7.7	1493	1	US-08-593-535-24	Sequence 24, Appl
39	54.2	7.7	289	1	US-08-341-568-3	Sequence 3, Appli
40	54.2	7.7	289	2	US-08-911-020-3	Sequence 3, Appli
41	54.2	7.7	1882	3	US-09-370-253-1	Sequence 1, Appli
42	54	7.7	991	3	US-08-924-747-25	Sequence 25, Appl
43	54	7.7	1721	4	PCT-US96-00994-3	Sequence 4, Appli
44	54	7.7	2550	5	5258287-23	Patent No. 5258287
45	53.8	7.6	1700	2	US-08-897-340-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-08-909-965C-9
; Sequence 9, Application US/08909965C
; Patent No. 5936078
; GENERAL INFORMATION:
;   APPLICANT: Kuga Tetsro
;   APPLICANT: Nakagawa Satoshi
;   APPLICANT: Sakaki yoshiyuki
;   APPLICANT: Zhao Nanding
;   APPLICANT: Hashida Hideji
;   TITLE OF INVENTION: NOVEL DNA, NOVEL POLYPEPTIDE
;   TITLE OF INVENTION: AND NOVEL ANTIBODY
;   NUMBER OF SEQUENCES: 17
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: FITZPATRICK, CELLA, HARPER AND SCINTO
;     STREET: 277 Park Avenue
;     CITY: New York
;     STATE: New York
;     COUNTRY: U.S.A.
;     ZIP: 10172-0194
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/909,965C
;     FILING DATE: August 12, 1997
;     CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 322745/95
;     APPLICATION NUMBER: PCT/JP96/03630
;     FILING DATE: 12-No. 5936078-1995
;     FILING DATE: 12-Dec-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Lawrence S. Perry
;     REGISTRATION NUMBER: 31865
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 212-758-2400
;     TELEFAX: 212-758-2982
;     TELEX: 236262
; INFORMATION FOR SEQ ID NO: 9:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 882 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA to mRNA
;   ORIGINAL SOURCE:
;     ORGANISM: human
;   IMMEDIATE SOURCE:
```

```
;      CLONE:  F1180
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  97 to 195
;      IDENTIFICATION METHOD:  by experiment
US-08-909-965C-9


  Query Match             8.7%;  Score 61.2;  DB 2;  Length 882;
  Best Local Similarity   76.5%;  Pred. No. 1.8e-05;
  Matches   75;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy     608 ctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaa 667
           || |   || |  ||||  |||  |  |    |||||| ||| |||  || | |||||||
Db     785 CTTTCAATTTATAAATTTGTTATCTTAATAAATATTATAAAAATGAAAAAAAAAAAAAAA 844

Qy     668 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||||
Db     845 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 882


RESULT   2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             8.5%;  Score 59.6;  DB 1;  Length 7218;
  Best Local Similarity   7.0%;  Pred. No. 8.2e-05;
  Matches   30;  Conservative  225;  Mismatches  175;  Indels    0;  Gaps    0;

Qy     225 tccccgtcggctgcacctgcgtgctgccccgttcagtgtgaccgccaaggccgtggggcc 284
           | |  ||||   |  | ||||  :: :::: ::: : :  :: ::    :: :    ::
Db    1044 TGCAGGTCGAGGGAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1103

Qy     285 cttagactggacacgtgtgctccccagagggcacccccctatttatgtgtatttattgtta 344
           :::   ::   : : : : ::::::      : :::::: ::: : : : ::: :: ::
Db    1104 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1163

Qy     345 tttatatgcctcccccaacactacccttggggtctgggcattccccgtgtctggaggaca 404
           ::: : : ::::::::  : :: :::::    :::   : ::::::: : :::      :
Db    1164 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1223

Qy     405 gccccccactgttctcctcatctccagcctcagtagttgggggtwgaaggagctcagcac 464
           :::::: :: :::::::: :::::  ::::  : ::    ::       ::   : : :
Db    1224 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1283

Qy     465 ctcttccagcccttaaagctgcagaaaaggtgtcacacggctgcctgtaccttggytccc 524
           ::::::: :::::    :: :        : :: : :  :: ::: : :::: |::::
Db    1284 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1343

Qy     525 tgtcctgctcccggcttccctaccctatcactggcctcaggcccccgcaggctgcctct 584
           : :::: :::::  :::::::: :::: :: ::  ::::   ::::: :   :: :::::
Db    1344 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1403

Qy     585 tcccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtatta 644
           ::::  :::::::     : ::::: :::::  | ||| |  ||   | |   |   |
Db    1404 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCAAATTCTTCTATCTCTTTAACTAC 1463

Qy     645 ttaaactgat 654
           ||  |  |||
Db    1464 TTGCATAGAT 1473


RESULT   3
US-08-530-950-3
; Sequence 3, Application US/08530950
; Patent No. 5736381
;  GENERAL INFORMATION:
;    APPLICANT:  Davis, Roger J.
;    APPLICANT:  Raingeaud, Joel
;    APPLICANT:  Gupta, Shashi
;    APPLICANT:  Derijard, Benoit
;    TITLE OF INVENTION:  CYTOKINE-, STRESS-, AND
;    TITLE OF INVENTION:  ONCOPROTEIN-ACTIVATED HUMAN PROTEIN KINASE
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/530,950
;      FILING DATE:  19-SEP-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. Peter
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  07917/010001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  3:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1602 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-530-950-3


  Query Match             8.3%;  Score 58.2;  DB 1;  Length 1602;
  Best Local Similarity   70.3%;  Pred. No. 0.0001;
  Matches   78;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy     595 cttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgat 654
           ||| | | | |    | |  ||||| | ||| | | ||||      | |  | ||   |
Db    1467 CTTAGGACTTCAAAAGGTGATTAAATATTTAATGATGTGTCATATGAGTCCTCAAAAAAA 1526

Qy     655 gaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           || | ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1527 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1577


RESULT   4
US-08-888-429A-3
; Sequence 3, Application US/08888429A
; Patent No. 6136596
;  GENERAL INFORMATION:
;    APPLICANT:  Davis, Roger J.
;    APPLICANT:  Whitmarsh, Alan
;    APPLICANT:  Tournier, Cathy
;    TITLE OF INVENTION:  CYTOKINE-, STRESS-, AND ONCOPROTEIN-
;    TITLE OF INVENTION:  ACTIVATED HUMAN PROTEIN KINASE KINASES
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/888,429A
;      FILING DATE:  07-JUL-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/530,950
;      FILING DATE:  19-SEP-1995
;      APPLICATION NUMBER:  08/446,083
;      FILING DATE:  19-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, Peter J.
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  07917/053001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  299354
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1602 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  244...1245
US-08-888-429A-3
```

```
  Query Match             8.3%;  Score 58.2;  DB 3;  Length 1602;
  Best Local Similarity   70.3%;  Pred. No. 0.0001;
  Matches   78;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy     595 cttggaagtacccctgtttcttaaacaattatttaagtgtacgtgtattattaaactgat 654
           ||| | | | |    | |  ||||| | ||| | | ||||      | |  | ||   |
Db    1467 CTTAGGACTTCAAAAGGTGATTAAATATTTAATGATGTGTCATATGAGTCCTCAAAAAAA 1526

Qy     655 gaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           || | ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1527 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1577


RESULT   5
US-08-184-632-1
; Sequence 1, Application US/08184632
; Patent No. 5646009
;  GENERAL INFORMATION:
;    APPLICANT:  RHOADS, ROBERT E
;    APPLICANT:  DEBENEDETTI, ARRICO
;    TITLE OF INVENTION:  HYBRID VECTOR AND METHOD RESULTING IN
;    TITLE OF INVENTION:  PROTEIN OVERPRODUCTION BY EUKARYOTIC CELLS
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lowe, Price, LeBlanc & Becker
;      STREET:  99 Canal Center Plaza, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/184,632
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/956,937
;      FILING DATE:  05-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MILLS, DEMETRA J
;      REGISTRATION NUMBER:  34,506
;      REFERENCE/DOCKET NUMBER:  434-025B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-684-1111
;      TELEFAX:  703-684-1124
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1897 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-184-632-1


  Query Match             8.3%;  Score 58.2;  DB 1;  Length 1897;
  Best Local Similarity   75.8%;  Pred. No. 0.00011;
  Matches   72;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy     611 tttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaa 670
           ||  |  ||   ||| | || || |   ||| ||| | |   ||| | | ||||||||||
Db    1797 TTATTACAATTTTTAATAAATTGAATAGGTAGTATCATATATATGGAAAAAAAAAAAAAA 1856

Qy     671 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||
Db    1857 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1891
```

```
RESULT   6
US-08-545-196B-10
; Sequence 10, Application US/08545196B
; Patent No. 6080577
;  GENERAL INFORMATION:
;    APPLICANT:  MELKI, JUDITH
;    APPLICANT:  MUNNICH, ARNOLD
;    TITLE OF INVENTION:  SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
;    TITLE OF INVENTION:  FOR SPINAL MUSCULAR ATROPHY
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH, LLP
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,196B
;      FILING DATE:  19-OCT-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FARACI, C. J.
;      REGISTRATION NUMBER:  32,350
;      REFERENCE/DOCKET NUMBER:  2121-110P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1582 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-545-196B-10


  Query Match             8.2%;  Score 57.6;  DB 3;  Length 1582;
  Best Local Similarity   75.0%;  Pred. No. 0.00014;
  Matches   72;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy     610 gtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaa 669
           || | | | | || |||||||| | |    |  |  |  |||    |  || | |||||||||
Db    1461 GTATATAATAAAAATATTTAATTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1520

Qy     670 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||| ||||||||||||||
Db    1521 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1556


RESULT   7
US-08-545-196B-12
; Sequence 12, Application US/08545196B
; Patent No. 6080577
;  GENERAL INFORMATION:
;    APPLICANT:  MELKI, JUDITH
;    APPLICANT:  MUNNICH, ARNOLD
;    TITLE OF INVENTION:  SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
;    TITLE OF INVENTION:  FOR SPINAL MUSCULAR ATROPHY
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH, LLP
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,196B
;      FILING DATE:  19-OCT-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FARACI, C. J.
;      REGISTRATION NUMBER:  32,350
;      REFERENCE/DOCKET NUMBER:  2121-110P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1582 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-545-196B-12


  Query Match             8.2%;  Score 57.6;  DB 3;  Length 1582;
  Best Local Similarity   75.0%;  Pred. No. 0.00014;
  Matches   72;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy     610 gtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaa 669
           || | | | | || |||||||| | |    |  |  |  |||    |  || | |||||||||
Db    1461 GTATATAATAAAAATATTTAATTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1520

Qy     670 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||| ||||||||||||||
Db    1521 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1556


RESULT   8
US-08-841-349-18/c
; Sequence 18, Application US/08841349B
; Patent No. 5955594
; GENERAL INFORMATION:
;  APPLICANT: MISHRA, LOPA
;  TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT. . .
;  FILE REFERENCE: XX/PO4470US0
;  CURRENT APPLICATION NUMBER: US/08/841,349B
;  CURRENT FILING DATE: 1997-04-30
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 467
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   OTHER INFORMATION: For all n's in this sequence, n=(a or g or c or t)
US-08-841-349-18


  Query Match             8.1%;  Score 57.4;  DB 2;  Length 467;
  Best Local Similarity   58.6%;  Pred. No. 0.0001;
  Matches   85;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy     558 ggcctcaggcccccgcaggctgcctcttcccaacctccttggaagtacccctgtttctta 617
           ||   |||||    ||   || |      |     || ||||||||          |    |
Db     152 GGGNCCAGGCTNTGGCTNTCTTCTAACCCNTCCCCCCCCTTGGAAANNTTTTNATAAAAAA 93

Qy     618 aacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaaaaaaaaa 677
           || ||   |||||   |   |        |||   |  || | |||||||||||||||||
Db      92 AAAAAAACTTTAAAANTTTNTTNNNNNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 33
```

```
Qy    678 aaaaaaaaaaaaaaaaaaaaaaa 702
          |||||||||||||||||||||||
Db     32 AAAAAAAAAAAAAAAAAAAAAAA 8


RESULT   9
5378464-1
;Patent No. 5378464
;     APPLICANT: MCEVER, RODGER P.
;     TITLE OF INVENTION: MODULATION OF INFLAMMATORY RESPONSES
;BY ADMINISTRATION OF GMP-140 OR ANTIBODY TO GMP-140
;     NUMBER OF SEQUENCES: 32
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/320,408
;       FILING DATE: 08-MAR-1989
;SEQ ID NO:1:
;       LENGTH: 2989
5378464-1


  Query Match             8.1%;  Score 57;  DB 5;  Length 2989;
  Best Local Similarity   65.1%;  Pred. No. 0.00024;
  Matches   84;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy    577 ctgcctcttcccaacctccttggaagtacccctgtttcttaaacaattatttaagtgtac 636
          ||||| |    ||   ||| |  ||  |||   |     ||  ||| |   ||| |  |
Db   2844 ctgccaccagtcactgtccctcaaatgaccaaagaccaatattcaaatgcgtaattaaaa 2903

Qy    637 gtgtattattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 696
          |  | ||    ||   |  || | ||||||||||||||||||||||||||||||||||||
Db   2904 gaattttccccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2963

Qy    697 aaaaaaaaa 705
          |||||||||
Db   2964 aaaaaaaaa 2972


RESULT  10
US-09-248-335-35
; Sequence 35, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 35
;   LENGTH: 1098
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-35


  Query Match             8.0%;  Score 56.6;  DB 3;  Length 1098;
  Best Local Similarity   74.7%;  Pred. No. 0.00021;
  Matches   71;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy    611 tttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaa 670
          ||| ||      || |||||    |  | |||| |||   | |  || | ||||||||||
Db    968 tttattgtttggttctttaacaagagttatatttttaccatcaaaaaaaaaaaaaaaaa 1027

Qy    671 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
          |||||||||||||||||||||||||||||||||||
Db   1028 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1062


RESULT  11
US-08-361-467B-3
; Sequence 3, Application US/08361467B
; Patent No. 5633441
;  GENERAL INFORMATION:
;    APPLICANT:  De Greef, Willy
;    APPLICANT:  Van Emmelo, John
;    APPLICANT:  De Oliveria, Dulce E.
;    APPLICANT:  De Souza, Maria-Helena
;    APPLICANT:  Van Montagu, Marc
;    TITLE OF INVENTION:  PLANTS WITH MODIFIED FLOWERS, SEEDS OR
;    TITLE OF INVENTION:  EMBRYOS
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/361,467B
;      FILING DATE:  22-DEC-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/681,492
;      FILING DATE:  04-APR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/EP90/01275
;      FILING DATE:  01-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 89 402 224.3
;      FILING DATE:  04-AUG-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Schulman, Robert M.
;      REGISTRATION NUMBER:  31,196
;      REFERENCE/DOCKET NUMBER:  010830-027
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  748 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    IMMEDIATE SOURCE:
;      CLONE:  4B12
US-08-361-467B-3


  Query Match             8.0%;  Score 56.4;  DB 1;  Length 748;
  Best Local Similarity   80.5%;  Pred. No. 0.0002;
  Matches   66;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    624 tatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaa 683
          ||||||||   || || | ||   ||   |  || | ||||||||||||||||||||||
Db    663 TATTTAAGGTGACTTGGAATACATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 722

Qy    684 aaaaaaaaaaaaaaaaaaaaa 705
          |||||||||||||||||||||
Db    723 AAAAAAAAAAAAAAAAAAAAA 744


RESULT  12
US-08-484-332C-3
; Sequence 3, Application US/08484332C
; Patent No. 5767374
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  De Greef, Willy
;    APPLICANT:  Van Emmelo, John
;    APPLICANT:  De Oliveria, Dulce E.
;    APPLICANT:  De Souza, Maria-Helena
;    APPLICANT:  Van Montagu, Marc
;    TITLE OF INVENTION:  PLANTS WITH MODIFIED FLOWERS, SEEDS OR
;    TITLE OF INVENTION:  EMBRYOS
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,332C
;      FILING DATE:  7-JUNE-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/361,467
;      FILING DATE:  22-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/681,492
;      FILING DATE:  04-APR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/EP90/01275
;      FILING DATE:  01-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 89 402 224.3
;      FILING DATE:  04-AUG-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Schulman, Robert M.
;      REGISTRATION NUMBER:  31,196
;      REFERENCE/DOCKET NUMBER:  010830-093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  748 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    IMMEDIATE SOURCE:
;      CLONE:  4B12
US-08-484-332C-3


  Query Match             8.0%;  Score 56.4;  DB 1;  Length 748;
  Best Local Similarity   80.5%;  Pred. No. 0.0002;
  Matches   66;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy     624 tatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaaaaaaaaaaaaaaa 683
           ||||||||   || || | ||   ||   |  || | ||||||||||||||||||||||
Db     663 TATTTAAGGTGACTTGGAATACATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 722

Qy     684 aaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||
Db     723 AAAAAAAAAAAAAAAAAAAAA 744


RESULT  13
US-08-242-677-1
; Sequence 1, Application US/08242677
; Patent No. 5677143
;  GENERAL INFORMATION:
;    APPLICANT:  Gaynor, Richard B
;    APPLICANT:  Wu, Foon W.
;    TITLE OF INVENTION:  Cellular Nucleic Acid Binding Protein
;    TITLE OF INVENTION:  and Uses Thereof in regulating Gene Expression and in the
;    TITLE OF INVENTION:  Treatment of AIDS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/242,677
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  UTSD:401
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713-787-1400
;      TELEFAX:  713-789-2679
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5173 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4863
US-08-242-677-1


  Query Match             8.0%;  Score 56.4;  DB 1;  Length 5173;
  Best Local Similarity   73.5%;  Pred. No. 0.00039;
  Matches   72;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy     608 ctgtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaa 667
           ||   | | | | |||  ||||||     ||  | ||  |||   |  || | |||||||
Db    5068 CTTCCTTTAATAAAATATTTTAAGCAATTGTCCAATAAAAAAAAAAAAAAAAAAAAAAA 5127

Qy     668 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||
Db    5128 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5165


RESULT  14
US-08-265-087-3
; Sequence 3, Application US/08265087
; Patent No. 5571515
;  GENERAL INFORMATION:
;    APPLICANT:  Scott, Phillip
;    APPLICANT:  Trinchieri, Giorgio
;    TITLE OF INVENTION:  Compositions and Methods for Use of
;    TITLE OF INVENTION:  IL-12 as an Adjuvant
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Center, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/265,087
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/229,282
;      FILING DATE:  18-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST51AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9206
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1364 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  101..859
US-08-265-087-3


  Query Match             7.9%;  Score 56;  DB 1;  Length 1364;
  Best Local Similarity   74.0%;  Pred. No. 0.0003;
  Matches   71;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy     610 gtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaa 669
           || | ||  | || |||||    ||       |  || ||| |  ||||   |||||||||
Db    1267 GTGTATTTGAAAAATATTTTTAAGTGTTCTAAAAATAAAAGTATTGAATTAAAAAAAAAA 1326

Qy     670 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||
Db    1327 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1362


RESULT  15
US-08-621-493-3
; Sequence 3, Application US/08621493
; Patent No. 5723127
;  GENERAL INFORMATION:
;    APPLICANT:  Scott, Phillip
;    APPLICANT:  Trinchieri, Giorgio
;    TITLE OF INVENTION:  Compositions and Methods for Use of
;    TITLE OF INVENTION:  IL-12 as an Adjuvant
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Center, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,493
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/265,087
;      FILING DATE:  17-JUN-1994
;      APPLICATION NUMBER:  US 08/229,282
;      FILING DATE:  18-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST51AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9206
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1364 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  101..859
US-08-621-493-3


  Query Match             7.9%;  Score 56;  DB 1;  Length 1364;
  Best Local Similarity   74.0%;  Pred. No. 0.0003;
  Matches   71;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy     610 gtttcttaaacaattatttaagtgtacgtgtattattaaactgatgaacacaaaaaaaaa 669
           || | ||  | || |||||    ||       |  || ||| |  ||||   |||||||||
Db    1267 GTGTATTTGAAAAATATTTTTAAGTGTTCTAAAAATAAAAGTATTGAATTAAAAAAAAAA 1326

Qy     670 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
           |||||||||||||||||||||||||||||||||||||
Db    1327 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1362


Search completed: February 24, 2001, 21:06:11
Job time: 6969 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:47:55 ; Search time 26.41 Seconds
(without alignments)
112.641 Million cell updates/sec

Title: US-09-320-713-2
Perfect score: 463
Sequence: 1 ARVDTDEDRYPQKLAFAECL.....FHTEFIHVPVGCTCVLPRSV 87

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	463	100.0	87	21	Y53890	Partial amino acid
2	459	99.1	197	21	Y92238	Human interleukin-
3	459	99.1	197	21	Y44460	Human Interleukin
4	459	99.1	197	21	Y53892	Amino acid sequenc
5	459	99.1	206	21	Y44485	Human Interleukin
6	459	99.1	425	21	Y44462	Human Interleukin
7	143.5	31.0	160	21	Y53891	Partial amino acid
8	143.5	31.0	173	21	Y53893	Partial amino acid
9	143.5	31.0	185	21	Y70656	Mature human trans
10	143.5	31.0	186	21	Y70655	Mature human trans
11	143.5	31.0	187	21	Y70654	Mature human trans
12	143.5	31.0	187	21	Y70663	Mature human Ztgfb
13	143.5	31.0	202	21	Y70653	Human transforming
14	143.5	31.0	209	21	Y70662	Human transforming
15	136	29.4	153	18	W28514	Product of clone B
16	136	29.4	163	18	W13651	Human CTLA-8. Hom
17	136	29.4	163	21	Y81986	Human CTLA-8 prote
18	132.5	28.6	183	21	Y70658	Mature murine tran
19	132.5	28.6	205	21	Y70657	Murine transformin
20	125	27.0	79	16	R76572	Human CTLA-8 exon-
21	125	27.0	155	16	R76573	Human CTLA-8. Hom
22	124	26.8	151	16	R76571	Herpesvirus ORF13
23	124	26.8	151	17	W02387	HVS13 (viral homol
24	124	26.8	151	18	W13653	Herpesvirus Saimir
25	124	26.8	151	20	W92413	Herpesvirus Saimir
26	124	26.8	151	21	Y81988	Herpes virus CTLA-
27	121	26.1	132	16	R76574	Human CTLA-8 matur
28	119	25.7	150	16	R76570	Mouse CTLA-8. Mus
29	119	25.7	150	18	W13652	Rat CTLA-8. Rattu
30	119	25.7	150	21	Y81987	Rat CTLA-8 protein
31	119	25.7	158	17	W02386	Murine CTLA8 (inte
32	119	25.7	158	20	W92412	Murine CTLA-8 prot
33	110.5	23.9	158	20	W85646	Murine Zcyto7 matu
34	110.5	23.9	160	20	W85645	Murine Zcyto7 matu
35	110.5	23.9	180	20	Y28236	Mouse interleukin-
36	110.5	23.9	180	20	Y22198	Murine EDIRF I pro
37	110.5	23.9	180	20	W85617	Murine Zcyto7. Mu
38	105.5	22.8	408	21	Y44461	Human Interleukin
39	104.5	22.6	160	20	W85625	Human Zcyto7 matur
40	98.5	21.3	128	20	W85643	Human Zcyto7 matur
41	98.5	21.3	151	20	W85641	Human Zcyto7 matur
42	98.5	21.3	153	20	W85642	Human Zcyto7 matur
43	98.5	21.3	154	20	W85640	Human Zcyto7 matur
44	98.5	21.3	157	20	W85644	Human Zcyto7 matur
45	98.5	21.3	158	20	W85639	Human Zcyto7 matur

ALIGNMENTS

RESULT 1
Y53890
ID Y53890 standard; Protein; 87 AA.
XX
AC Y53890;
XX
DT 13-MAR-2000 (first entry)
XX
DE Partial amino acid sequence of human interleukin-21.
XX
KW Human; interleukin-21; IL-21; IL-22; immune system disorder;
KW immune cell chemotaxis; haematopoietic cell disorder;
KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW allergic asthma; respiratory problem; organ rejection;
KW graft-versus-host disease; GVHD; inflammation;
KW hyperproliferative disorder; tissue regeneration;
KW embryonic stem cell differentiation; embryonic stem cell proliferation;
KW haematopoietic lineage.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 3..11
FT /note= "conserved domain I"
FT Domain 19..24
FT /note= "conserved domain II"
FT Domain 46..52
FT /note= "conserved domain III"
FT Domain 75..82
FT /note= "conserved domain IV"
XX
PN WO9961617-A1.
XX
PD 02-DEC-1999.
XX

PF 27-MAY-1999; 99WO-US11644.
XX
PR 29-MAY-1998; 98US-0087340.
PR 10-SEP-1998; 98US-0099805.
PR 30-APR-1999; 99US-0131965.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ebner R;
XX
DR WPI; 2000-072622/06.
DR N-PSDB; Z36834.
XX
PT Novel polynucleotides used to develop products for treating e.g. immune
PT disorders, blood disorders, autoimmune disorders, allergies,
PT inflammation, hyperproliferative disorders or infections -
XX
PS Claim 25; Fig 1; 170pp; English.
XX
CC The present sequence represents a partial human interleukin-21 (IL-21)
CC protein. The specification also describes IL-22 polynucleotides and
CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC of apoptotic T-cells. IL-21 and IL-22 may be useful in treating
CC deficiencies or disorders of the immune system, by activating or
CC inhibiting the proliferation, differentiation, or mobilization
CC (chemotaxis) of immune cells, treating or detecting deficiencies or
CC disorders of haematopoietic cells, to modulate haemostatic or
CC thrombolytic activity, in treating or detecting autoimmune disorders,
CC treating asthma (particularly allergic asthma) or other respiratory
CC problems, to treat and/or prevent organ rejection or graft-versus-host
CC disease (GVHD), to modulate inflammation, to treat or detect
CC hyperproliferative disorders, to treat or detect infectious agents, to
CC differentiate, proliferate and attract cells, leading to the regeneration
CC of tissues, IL-21 and IL-22 may also increase or decrease the
CC differentiation or proliferation of embryonic stem cells and
CC haematopoietic lineage, may be used to modulate mammalian
CC characteristics.
XX
SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 21; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.9e-52;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGL 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 arvdtdedrypqklafaeclcrgcidartgretaalnsvrllqsllvlrrrpcsrdgsgl 60

Qy 61 PTPGAFAFHTEFIHVPVGCTCVLPRSV 87
|||||||||||||||||||||||||||
Db 61 ptpgafafhtefihvpvgctcvlprsv 87

RESULT 2
Y92238
ID Y92238 standard; Protein; 197 AA.
XX
AC Y92238;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human interleukin-17 (IL-17) homologue.
XX
KW Interleukin 17; IL-17; haematopoiesis; chemotherapy; cytostatic;
KW antianemic; cardiant; hemostatic; anti-inflammatory; anti-HIV.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= signal_peptide
FT Protein 19..197
FT /label= mature_protein
XX
PN WO200020593-A1.
XX
PD 13-APR-2000.
XX
PF 30-SEP-1999; 99WO-US22678.
XX
PR 02-OCT-1998; 98US-0102883.
PR 01-DEC-1998; 98US-0110405.
PR 11-JUN-1999; 99US-0138910.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Glasebrook AL, Su EW, Wei J, Liu L;
XX
DR WPI; 2000-303778/26.
DR N-PSDB; A09153.
XX
PT Nucleic acid encoding an interleukin-17 (IL-17) homolog polypeptide
PT which enhances hematopoiesis, useful for treating e.g. anemia,
PT thrombocytopenia, viral and bacterial infections
XX
PS Claim 16; Page 92-93; 111pp; English.
XX
CC Interleukin 17 (IL-17) stimulates hematopoiesis and production of
CC neutrophils, granulocytes, or platelets, this may be useful during
CC chemotherapy. IL-17 homologues have at least one activity selected
CC from induction of cytotoxic T cells, induction of lymphokine-activated
CC killer cell proliferation or a B or T cell stimulation. The IL-17
CC homologue may also be used to treat viral or bacterial infections,
CC immune related diseases, anemia, leukemia, thrombocytopenia, uremia,
CC Von Willebrand disease, postoperative cardiovascular dysfunction,
CC treatment of AIDS (acquired immune deficiency syndrome)-related bone
CC marrow failure, and inflammatory diseases of the gastrointestinal
CC system, joints, and lungs.
XX
SQ Sequence 197 AA;

Query Match 99.1%; Score 459; DB 21; Length 197;
Best Local Similarity 100.0%; Pred. No. 3.3e-51;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 112 rvdtdedrypqklafaeclcrgcidartgretaalnsvrllqsllvlrrrpcsrdgsglp 171

Qy 62 TPGAFAFHTEFIHVPVGCTCVLPRSV 87
|||||||||||||||||||||||||||
Db 172 tpgafafhtefihvpvgctcvlprsv 197

RESULT 3
Y44460
ID Y44460 standard; Protein; 197 AA.
XX
AC Y44460;
XX
DT 27-MAR-2000 (first entry)
XX
DE Human Interleukin 17C, PRO1122 polypeptide.
XX
KW Interleukin; IL-17C; PRO1122 polypeptide; clone DNA62377-1381-1; UNQ561;
KW cytokine IL-17; cytotoxic T-lymphocyte-associated antigen 8; CTLA-8;
KW hybridisation probe; antagonist; degenerative cartilaginous disorder;
KW agonist; diagnose; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= Signal_peptide

```
FT   Protein          19..197
FT                    /label= Mature_IL-17C_polypeptide
FT                    /note= "Used to treat degenerative cartilaginous
FT                    disorder"
FT   Misc-difference  109
FT                    /note= "Conserved Trp residue"
FT   Misc-difference  129
FT                    /note= "Conserved Cys residue"
FT   Misc-difference  134
FT                    /note= "Conserved Cys residue"
FT   Misc-difference  163
FT                    /note= "Conserved Cys residue"
FT   Misc-difference  189
FT                    /note= "Conserved Cys residue"
FT   Misc-difference  191
FT                    /note= "Conserved Cys residue"
XX
PN   WO9960127-A2.
XX
PD   25-NOV-1999.
XX
PF   14-MAY-1999;   99WO-US10733.
XX
PR   15-MAY-1998;   98US-0085579.
PR   23-DEC-1998;   98US-0113621.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Filvaroff E,  Goddard A,  Gurney AL,  Li H,  Wood WI;
XX
DR   WPI; 2000-116314/10.
DR   N-PSDB; Z29728.
XX
PT   New polypeptides designated PRO1031 and PRO1122 used to treat a
PT   degenerative cartilaginous disorder -
XX
PS   Claim 23; Fig 3; 141pp; English.
XX
CC   The present sequence is the human PRO1122 polypeptide, also referred to
CC   as UNQ561, and as interleukin-17C (IL-17C), encoded by
CC   clone DNA62377-1381-1. This sequence has identity with the
CC   cytokine IL-17 and cytotoxic T-lymphocyte-associated antigen 8 (CTLA-8)
CC   and has leucine zipper pattern. PRO1122 is expressed in pancreas, small
CC   intestine, stomach and testis also. It shares about 26-28% amino acid
CC   identity with IL-17 and IL-17B. The entire coding region of IL-17C can
CC   be used as hybridisation probe. The PRO1122 polypeptide, agonist or
CC   antagonist, is used to diagnose and treat a degenerative cartilaginous
CC   disorder.
XX
SQ   Sequence   197 AA;


  Query Match             99.1%;  Score 459;  DB 21;  Length 197;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-51;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 rvdtdedrypqklafaeclcrgcidartgretaalnsvrllqslvlrrrpcsrdgsglp 171

Qy      62 TPGAFAFHTEFIHVPVGCTCVLPRSV 87
           |||||||||||||||||||||||||
Db     172 tpgafafhtefihvpvgctcvlprsv 197


RESULT   4
Y53892
ID   Y53892 standard; Protein; 197 AA.
XX
AC   Y53892;
XX
DT   13-MAR-2000  (first entry)
XX
DE   Amino acid sequence of human interleukin-21.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /note= "signal peptide"
FT   Domain          34..40
FT                   /note= "conserved domain V"
FT   Domain          63..68
FT                   /note= "conserved domain VI"
FT   Domain          104..109
FT                   /note= "conserved domain VII"
FT   Domain          113..121
FT                   /note= "conserved domain I"
FT   Domain          129..134
FT                   /note= "conserved domain II"
FT   Domain          156..162
FT                   /note= "conserved domain III"
FT   Domain          185..192
FT                   /note= "conserved domain IV"
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
DR   N-PSDB; Z36836.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Claim 26; Fig 6A-B; 170pp; English.
XX
CC   The present sequence represents a human interleukin-21 (IL-21)
CC   protein. The specification also describes IL-22 polynucleotides and
CC   polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC   of apoptotic T-cells. IL-21 and IL-22 may be useful in treating
CC   deficiencies or disorders of the immune system, by activating or
CC   inhibiting the proliferation, differentiation, or mobilization
CC   (chemotaxis) of immune cells, treating or detecting deficiencies or
CC   disorders of haematopoietic cells, to modulate haemostatic or
CC   thrombolytic activity, in treating or detecting autoimmune disorders,
CC   treating asthma (particularly allergic asthma) or other respiratory
CC   problems, to treat and/or prevent organ rejection or graft-versus-host
CC   disease (GVHD), to modulate inflammation, to treat or detect
CC   hyperproliferative disorders, to treat or detect infectious agents, to
CC   differentiate, proliferate and attract cells, leading to the
CC   regeneration of tissues, IL-21 and IL-22 may also increase or decrease
CC   the differentiation or proliferation of embryonic stem cells and
CC   haematopoietic lineage, may be used to modulate mammalian
CC   characteristics.
XX
SQ   Sequence   197 AA;
```

```
  Query Match             99.1%;  Score 459;  DB 21;  Length 197;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-51;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 rvdtdedrypqklafaeclcrgcidartgretaalnsvrllqsllvlrrrpcsrdgsglp 171

Qy      62 TPGAFAFHTEFIHVPVGCTCVLPRSV 87
           |||||||||||||||||||||||||||
Db     172 tpgafafhtefihvpvgctcvlprsv 197


RESULT   5
Y44485
ID   Y44485 standard; Protein; 206 AA.
XX
AC   Y44485;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Human Interleukin 17C with C-terminal Gly(His)8 tag, IL-17C.his.
XX
KW   Interleukin; IL-17C.his; PRO1122 polypeptide; clone DNA62377-1381-1;
KW   immunoprecipitation; IL-17 receptor extracellular domain; IL-17R ECD;
KW   cytokine IL-17; hybridisation probe; antagonist; Gly(His)8 tag; agonist;
KW   degenerative cartilaginous disorder; diagnose; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /label= Signal_peptide
FT   Protein         19..197
FT                   /label= Mature_IL-17C_polypeptide
FT                   /note= "Used to treat degenerative cartilaginous
FT                   disorder"
FT   Misc-difference 109
FT                   /note= "Conserved Trp residue"
FT   Misc-difference 129
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 134
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 163
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 189
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 191
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 198..206
FT                   /note= "C-terminal Gly(His)8 tag"
XX
PN   WO9960127-A2.
XX
PD   25-NOV-1999.
XX
PF   14-MAY-1999;   99WO-US10733.
XX
PR   15-MAY-1998;   98US-0085579.
PR   23-DEC-1998;   98US-0113621.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Filvaroff E,  Goddard A,  Gurney AL,  Li H,  Wood WI;
XX
DR   WPI; 2000-116314/10.
XX
PT   New polypeptides designated PRO1031 and PRO1122 used to treat a
PT   degenerative cartilaginous disorder -
XX
PS   Example 11; Page 138-139; 141pp; English.
XX
CC   The present sequence is the human PRO1122 polypeptide, with a C-terminal
CC   Gly(His)8 tag, IL-17C.his, derived from the clone DNA62377-1381-1.
CC   This sequence is used in a competitive binding experiment for the
CC   immunoprecipitation of IL-17 receptor extracellular domain (IL-17R ECD).
CC   The entire coding region of IL-17C can be used as hybridisation probe.
CC   The PRO1122 polypeptide, agonist or antagonist, is used to diagnose and
CC   treat a degenerative cartilaginous disorder.
XX
SQ   Sequence   206 AA;


  Query Match             99.1%;  Score 459;  DB 21;  Length 206;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-51;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 rvdtdedrypqklafaeclcrgcidartgretaalnsvrllqsllvlrrrpcsrdgsglp 171

Qy      62 TPGAFAFHTEFIHVPVGCTCVLPRSV 87
           |||||||||||||||||||||||||||
Db     172 tpgafafhtefihvpvgctcvlprsv 197


RESULT   6
Y44462
ID   Y44462 standard; Protein; 425 AA.
XX
AC   Y44462;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Human Interleukin 17C-IgG1 Fc fusion protein, hIL-17C.fc.
XX
KW   Interleukin; IL-17C.fc; fusion protein; PRO1122 polypeptide; cytokine;
KW   human IgG1; fluoresence-activated cell sorter analysis; FACS;
KW   Tumour Necrosis Factor-alpha; TNF-alpha; leukemic monocyte; THP-1 cell.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /label= Signal_peptide
FT   Protein         19..197
FT                   /label= Mature_IL-17C_polypeptide
FT                   /note= "Used to treat degenerative cartilaginous
FT                   disorder"
FT   Misc-difference 109
FT                   /note= "Conserved Trp residue"
FT   Misc-difference 129
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 134
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 163
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 189
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 191
FT                   /note= "Conserved Cys residue"
FT   Region          197..425
FT                   /note= "Sequence derived from Fc region of human IgG1"
XX
PN   WO9960127-A2.
XX
PD   25-NOV-1999.
XX
PF   14-MAY-1999;   99WO-US10733.
XX
PR   15-MAY-1998;   98US-0085579.
PR   23-DEC-1998;   98US-0113621.
XX
PA   (GETH ) GENENTECH INC.
```

```
XX
PI   Chen J,  Filvaroff E,  Goddard A,  Gurney AL,  Li H,  Wood WI;
XX
DR   WPI; 2000-116314/10.
XX
PT   New polypeptides designated PRO1031 and PRO1122 used to treat a
PT   degenerative cartilaginous disorder -
XX
PS   Example 12; Page 129-130; 141pp; English.
XX
CC   The present sequence is the human IL-17C.fc fusion protein, derived from
CC   PRO1122 polypeptide and the Fc region of human IgG1. The cytokine IL-17C
CC   can be used to induce the release of TNF-alpha from human leukemic
CC   monocytic, THP-1 cells. The fusion protein, IL-17C.fc is used to identify
CC   the binding of IL-17C to THP-1 cells, using fluoresence-activated cell
CC   sorter analysis (FACS).
XX
SQ   Sequence   425 AA;


  Query Match             99.1%;  Score 459;  DB 21;  Length 425;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-51;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 rvdtdedrypqklafaeclcrgcidartgretaalnsvrllqslllvlrrrpcsrdgsglp 171

Qy      62 TPGAFAFHTEFIHVPVGCTCVLPRSV 87
           |||||||||||||||||||||||||||
Db     172 tpgafafhtefihvpvgctcvlprsv 197


RESULT   7
Y53891
ID   Y53891 standard; Protein; 160 AA.
XX
AC   Y53891;
XX
DT   13-MAR-2000  (first entry)
XX
DE   Partial amino acid sequence of human interleukin-22.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   26
FT                   /note= "potential N-linked glycosylation site"
FT   Domain          57..64
FT                   /note= "conserved domain I"
FT   Domain          72..77
FT                   /note= "conserved domain II"
FT   Domain          99..105
FT                   /note= "conserved domain III"
FT   Domain          121..128
FT                   /note= "conserved domain IV"
FT   Modified-site   139
FT                   /note= "potential N-linked glycosylation site"
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
DR   N-PSDB; Z36835.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Claim 27; Fig 2A-B; 170pp; English.
XX
CC   The present sequence represents a partial human interleukin-22 (IL-22)
CC   protein. The specification also describes IL-21 polynucleotides and
CC   polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC   of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating
CC   deficiencies or disorders of the immune system, by activating or
CC   inhibiting the proliferation, differentiation, or mobilization
CC   (chemotaxis) of immune cells, treating or detecting deficiencies or
CC   disorders of haematopoietic cells, to modulate haemostatic or
CC   thrombolytic activity, in treating or detecting autoimmune disorders,
CC   treating asthma (particularly allergic asthma) or other respiratory
CC   problems, to treat and/or prevent organ rejection or graft-versus-host
CC   disease (GVHD), to modulate inflammation, to treat or detect
CC   hyperproliferative disorders, to treat or detect infectious agents, to
CC   differentiate, proliferate and attract cells, leading to the
CC   regeneration of tissues, IL-21 and IL-22 may also increase or decrease
CC   the differentiation or proliferation of embryonic stem cells and
CC   haematopoietic lineage, may be used to modulate mammalian
CC   characteristics.
XX
SQ   Sequence   160 AA;


  Query Match             31.0%;  Score 143.5;  DB 21;  Length 160;
  Best Local Similarity   37.8%;  Pred. No. 7.7e-11;
  Matches   31;  Conservative   10;  Mismatches   32;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |:  |  |||: |  | ||||||:    | |     |  :    :|||| |    |  :
Db      55 risydparyprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsv- 113

Qy      62 TPGAFAFHTE-FIHVPVGCTCV 82
                 :|| :: :|||||||
Db     114 -------yteayvtipvgctcv 128


RESULT   8
Y53893
ID   Y53893 standard; Protein; 173 AA.
XX
AC   Y53893;
XX
DT   13-MAR-2000  (first entry)
XX
DE   Partial amino acid sequence of human interleukin-22.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
```

```
FT   Domain          18..23
FT                   /note= "conserved domain VI"
FT   Modified-site   39..41
FT                   /note= "Asn39 is a potential N-linked glycosylation site"
FT   Domain          60..65
FT                   /note= "conserved domain VII"
FT   Domain          69..77
FT                   /note= "conserved domain I"
FT   Domain          85..90
FT                   /note= "conserved domain II"
FT   Domain          112..118
FT                   /note= "conserved domain III"
FT   Domain          134..141
FT                   /note= "conserved domain IV"
FT   Modified-site   152..154
FT                   /note= "Asn152 is a potential N-linked glycosylation
FT                          site"
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
DR   N-PSDB; Z36837.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Disclosure; Fig 8; 170pp; English.
XX
CC   The present sequence represents a partial human interleukin-22 (IL-22)
CC   protein. The specification also describes IL-21 polynucleotides and
CC   polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC   of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating
CC   deficiencies or disorders of the immune system, by activating or
CC   inhibiting the proliferation, differentiation, or mobilization
CC   (chemotaxis) of immune cells, treating or detecting deficiencies
CC   of haematopoietic cells, to modulate haemostatic or thrombolytic
CC   activity, in treating or detecting autoimmune disorders, treating
CC   asthma (particularly allergic asthma) or other respiratory problems,
CC   to treat and/or prevent organ rejection or graft-versus-host disease
CC   (GVHD), to modulate inflammation, to treat or detect hyperproliferative
CC   disorders, to treat or detect infectious agents, to differentiate,
CC   proliferate and attract cells, leading to the regeneration of tissues,
CC   IL-21 and IL-22 may also increase or decrease the differentiation or
CC   proliferation of embryonic stem cells and haematopoietic lineage, may
CC   be used to modulate mammalian characteristics.
XX
SQ   Sequence   173 AA;
```

```
  Query Match             31.0%;  Score 143.5;  DB 21;  Length 173;
  Best Local Similarity   37.8%;  Pred. No. 8.4e-11;
  Matches   31;  Conservative   10;  Mismatches   32;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |:  |  |||: |  | ||| ||:    | |     |  :    :|||| |    |  :
Db      68 risydparyprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsv- 126

Qy      62 TPGAFAFHTE-FIHVPVGCTCV 82
                  :|| :: :|||||| |
Db     127 -------yteayvtipvgctcv 141
```

```
RESULT   9
Y70656
ID   Y70656 standard; Protein; 185 AA.
XX
AC   Y70656;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature human transforming growth factor beta-9, Ztgfbeta-9 protein-3.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52195.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 85-86; 97pp; English.
XX
CC   The present sequence is the mature human transforming growth factor
CC   beta-9, designated Ztgf beta-9. This is a mature sequence excluding the
CC   signal sequence extending from amino acid 18 to and including amino acid
CC   202 of Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   185 AA;
```

```
  Query Match             31.0%;  Score 143.5;  DB 21;  Length 185;
  Best Local Similarity   37.8%;  Pred. No. 9.1e-11;
  Matches   31;  Conservative   10;  Mismatches   32;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |:  |  |||: |  | ||||||:    | |     |  :    :|||| |    |  :
Db      80 risydparyprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsv- 138

Qy      62 TPGAFAFHTE-FIHVPVGCTCV 82
                  :|| :: :|||||||
Db     139 -------yteayvtipvgctcv 153
```

```
RESULT  10
Y70655
ID   Y70655 standard; Protein; 186 AA.
```

```
XX
AC   Y70655;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature human transforming growth factor beta-9, Ztgfbeta-9 protein-2.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52195.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 84; 97pp; English.
XX
CC   The present sequence is the mature human transforming growth factor
CC   beta-9, designated Ztgf beta-9. This is a mature sequence excluding the
CC   signal sequence extending from amino acid 17 to and including amino acid
CC   202 of Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   186 AA;


  Query Match             31.0%;  Score 143.5;  DB 21;  Length 186;
  Best Local Similarity   37.8%;  Pred. No. 9.1e-11;
  Matches   31;  Conservative   10;  Mismatches   32;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |:  |  |||: |  | |||||||:    | |     |  :    :|||| |    |  :
Db      81 risydparyprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsv- 139

Qy      62 TPGAFAFHTE-FIHVPVGCTCV 82
                  :|| :: :|||||||
Db     140 -------yteayvtipvgctcv 154


RESULT  11
Y70654
ID   Y70654 standard; Protein; 187 AA.
XX
AC   Y70654;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature human transforming growth factor beta-9, Ztgfbeta-9 protein-1.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52195.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 84; 97pp; English.
XX
CC   The present sequence is the mature human transforming growth factor
CC   beta-9, designated Ztgf beta-9. This is a mature sequence excluding the
CC   signal sequence extending from amino acid 16 to and including amino acid
CC   202 of Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   187 AA;


  Query Match             31.0%;  Score 143.5;  DB 21;  Length 187;
  Best Local Similarity   37.8%;  Pred. No. 9.2e-11;
  Matches   31;  Conservative   10;  Mismatches   32;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |:  |  |||: |  | |||||||:    | |     |  :    :|||| |    |  :
Db      82 risydparyprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsv- 140

Qy      62 TPGAFAFHTE-FIHVPVGCTCV 82
                  :|| :: :|||||||
Db     141 -------yteayvtipvgctcv 155


RESULT  12
Y70663
ID   Y70663 standard; Protein; 187 AA.
XX
AC   Y70663;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature human Ztgfbeta-9 variant protein.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
```

```
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52201.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 93-94; 97pp; English.
XX
CC   The present sequence is a mature variant human transforming growth
CC   factor beta-9 protein, designated Ztgf beta-9. Human Ztgf beta-9 was
CC   isolated from an arrayed pituitary gland cDNA plasmid library by PCR
CC   screening. This can be used to treat a variety of neurodegenerative
CC   diseases such as amyotrophic lateral sclerosis (ALS), Alzheimer's
CC   disease, Huntington's disease, Parkinson's disease and peripheral
CC   neuropathies, or demyelinating diseases including multiple sclerosis.
CC   Ztgf beta-9 peptides have antiviral activity and may also be used to
CC   regulate the proliferation, differentiation and apoptosis of neurons,
CC   glial cells, lymphocytes, hematopoietic cells and stromal cells.
XX
SQ   Sequence   187 AA;


  Query Match             31.0%;  Score 143.5;  DB 21;  Length 187;
  Best Local Similarity   37.8%;  Pred. No. 9.2e-11;
  Matches   31;  Conservative   10;  Mismatches   32;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |:  |  |||: |  | ||||||:    | |     |  :   :|||| |    |  :
Db      82 risydparyprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsv- 140

Qy      62 TPGAFAFHTE-FIHVPVGCTCV 82
                 :|| :: :|||||||
Db     141 -------yteayvtipvgctcv 155


RESULT  13
Y70653
ID   Y70653 standard; Protein; 202 AA.
XX
AC   Y70653;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human transforming growth factor beta-9, Ztgfbeta-9 protein.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..15
FT                   /label= Signal_peptide
FT   Protein         16..202
FT                   /note= "Mature transforming growth factor beta-9"
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52195.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 83; 97pp; English.
XX
CC   The present sequence is the human transforming growth factor beta-9,
CC   designated Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   202 AA;


  Query Match             31.0%;  Score 143.5;  DB 21;  Length 202;
  Best Local Similarity   37.8%;  Pred. No. 1e-10;
  Matches   31;  Conservative   10;  Mismatches   32;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |:  |  |||: |  | ||||||:    | |     |  :   :|||| |    |  :
Db      97 risydparyprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsv- 155

Qy      62 TPGAFAFHTE-FIHVPVGCTCV 82
                 :|| :: :|||||||
Db     156 -------yteayvtipvgctcv 170


RESULT  14
Y70662
ID   Y70662 standard; Protein; 209 AA.
XX
AC   Y70662;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human transforming growth factor beta-9, Ztgfbeta-9 variant protein.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /label= Signal_peptide
```

XX
PN WO200015798-A2.
XX
PD 23-MAR-2000.
XX
PF 17-SEP-1999; 99WO-US21677.
XX
PR 17-SEP-1998; 98US-0154817.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Presnell SR, Taft DW, Foley KP;
XX
DR WPI; 2000-271436/23.
DR N-PSDB; Z52201.
XX
PT Polynucleotides encoding a novel transforming growth factor beta-9
PT polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT antiproliferative agent -
XX
PS Claim 6; Page 92-93; 97pp; English.
XX
CC The present sequence is a variant of human transforming growth
CC factor beta-9, designated Ztgf beta-9. Human Ztgf beta-9 was isolated
CC from an arrayed pituitary gland cDNA plasmid library by PCR screening.
CC This can be used to treat a variety of neurodegenerative diseases such
CC as amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's
CC disease, Parkinson's disease and peripheral neuropathies, or
CC demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides
CC have antiviral activity and may also be used to regulate the
CC proliferation, differentiation and apoptosis of neurons, glial cells,
CC lymphocytes, hematopoietic cells and stromal cells.
XX
SQ Sequence 209 AA;

Query Match 31.0%; Score 143.5; DB 21; Length 209;
Best Local Similarity 37.8%; Pred. No. 1e-10;
Matches 31; Conservative 10; Mismatches 32; Indels 9; Gaps 2;

```
Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           |:  |  |||: |  | ||||||:    | |     |  :    :|||| |    |  :
Db     104 risydparyprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsv- 162

Qy      62 TPGAFAFHTE-FIHVPVGCTCV 82
                  :|| :: :|||||||
Db     163 -------yteayvtipvgctcv 177
```

RESULT 15
W28514
ID W28514 standard; Protein; 153 AA.
XX
AC W28514;
XX
DT 29-DEC-1997 (first entry)
XX
DE Product of clone B18.
XX
KW J5; J422; L105; H174-10; H174-43; B18; cytokine; PBMC;
KW peripheral blood mononuclear cell; disintegrin; metallo-protein;
KW Drosophila; leucine-rich repeat; monocyte; chemoattractant;
KW IP-10; CRG-2; CTLA-8; herpesvirus; Saimiri.
XX
OS Homo sapiens.
XX
PN WO9707198-A2.
XX
PD 27-FEB-1997.
XX
PF 08-AUG-1996; 96WO-US12897.
XX
PR 08-AUG-1996; 96WO-US12897.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Carlin M, Jacobs K, Kelleher K, Mccoy JM;
XX
DR WPI; 1997-165283/15.
DR N-PSDB; T87432.
XX
PT Polynucleotide(s) encoding proteins for treating, preventing and
PT ameliorating medical conditions - obtained from human activated
PT peripheral blood mononuclear cell, and murine adult thymus libraries
XX
PS Claim 51; Page 49-50; 61pp; English.
XX
CC This sequence was isolated from a human activated PBMC library using
CC a trap selecting for nucleotides encoding secreted proteins, and
CC encodes a protein having homology to murine CTLA-8 and herpesvirus
CC Saimiri ORF13.
XX
SQ Sequence 153 AA;

Query Match 29.4%; Score 136; DB 18; Length 153;
Best Local Similarity 39.3%; Pred. No. 6.7e-10;
Matches 33; Conservative 14; Mismatches 23; Indels 14; Gaps 3;

```
Qy       3 VDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRR--PCSRDGSGL 60
           |  | :||| ::  |:|   |||:|: |:|  ::||| : |  ||:||:   ||
Db      76 vtwdpnrypsevvqaqcrnlgcinaq-gkedismnsvpiqqetlvvrrkhqgcs------ 128

Qy      61 PTPGAFAFHTEFIHVPVGCTCVLP 84
                :|  | : | ||||||  |
Db     129 -----vsfqlekvlvtvgctcvtp 147
```

Search completed: February 16, 2001, 12:49:48
Job time: 113 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:49:25 ; Search time 38.3 Seconds
(without alignments)
266.243 Million cell updates/sec

Title: US-09-320-713-2
Perfect score: 463
Sequence: 1 ARVDTDEDRYPQKLAFAECL.....FHTEFIHVPVGCTCVLPRSV 87

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	459	99.1	197	4	Q9P0M4	Q9p0m4 homo sapien
2	136	29.4	78	4	Q9NUE6	Q9nue6 homo sapien
3	123	26.6	151	12	O40633	O40633 saimiriine
4	110.5	23.9	180	11	Q9QXT6	Q9qxt6 mus musculu
5	98.5	21.3	180	4	Q9UHF5	Q9uhf5 homo sapien
6	88	19.0	148	5	Q19778	Q19778 caenorhabdi
7	69.5	15.0	39	6	Q9TUP5	Q9tup5 bos taurus
8	69.5	15.0	39	6	Q9TUP4	Q9tup4 ovis aries
9	69.5	15.0	39	6	Q9TUP3	Q9tup3 sus scrofa
10	67	14.5	146	2	Q9S2C5	Q9s2c5 streptomyce
11	65	14.0	518	13	O13035	O13035 g proactiva
12	63.5	13.7	563	10	O24596	O24596 zea mays (m
13	62.5	13.5	745	5	Q9U195	Q9u195 leishmania
14	62	13.4	1737	4	O75097	O75097 homo sapien
15	61	13.2	874	11	Q9QYP0	Q9qyp0 rattus norv
16	61	13.2	2221	5	Q9U1M1	Q9u1m1 drosophila
17	61	13.2	2222	5	O97394	O97394 drosophila
18	60.5	13.1	315	5	Q9VUV7	Q9vuv7 drosophila
19	60	13.0	227	4	Q15935	Q15935 homo sapien
20	60	13.0	245	5	O45998	O45998 caenorhabdi
21	60	13.0	436	2	Q9XA68	Q9xa68 streptomyce
22	60	13.0	812	5	Q9VT12	Q9vt12 drosophila
23	59.5	12.9	270	13	O93391	O93391 coturnix co
24	59.5	12.9	491	5	Q18351	Q18351 caenorhabdi
25	59.5	12.9	544	5	Q20780	Q20780 caenorhabdi
26	59.5	12.9	581	10	O65478	O65478 arabidopsis
27	59.5	12.9	743	12	Q83111	Q83111 mouse adeno
28	59.5	12.9	938	10	Q9LDD7	Q9ldd7 lycopersico
29	59	12.7	198	11	Q9JJE6	Q9jje6 mus musculu
30	59	12.7	230	4	Q9NSC9	Q9nsc9 homo sapien
31	59	12.7	955	4	Q99466	Q99466 homo sapien
32	59	12.7	1999	4	Q99940	Q99940 homo sapien
33	59	12.7	2003	4	O00306	O00306 homo sapien
34	58.5	12.6	93	2	Q9RIS4	Q9ris4 streptomyce
35	58.5	12.6	102	2	Q9X8Y1	Q9x8y1 streptomyce
36	58.5	12.6	102	2	Q9L1X5	Q9l1x5 streptomyce
37	58.5	12.6	838	12	O41012	O41012 porcine ade
38	58.5	12.6	838	12	Q9YTR7	Q9ytr7 porcine ade
39	58.5	12.6	907	10	O64828	O64828 arabidopsis
40	58.5	12.6	2473	11	Q9QZ84	Q9qz84 mus musculu
41	58	12.5	183	2	Q9RZ51	Q9rz51 deinococcus
42	58	12.5	194	10	O22272	O22272 arabidopsis
43	58	12.5	347	4	Q9Y6U7	Q9y6u7 homo sapien
44	58	12.5	524	2	Q9RYT8	Q9ryt8 deinococcus
45	58	12.5	709	12	Q9JE95	Q9je95 trichomonas

ALIGNMENTS

RESULT 1
Q9P0M4
ID Q9P0M4 PRELIMINARY; PRT; 197 AA.
AC Q9P0M4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE INTERLEUKIN 17C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li H., Chen J., Huang A., Stinson J., Heldens S., Foster J., Dowd P.,
RA Gurney A.L., Wood W.I.;
RT "Cloning and Characterization of IL-17B and IL17C, Two New Members of
RT the IL-17 Cytokine Family.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:773-778(2000).
DR EMBL; AF152099; AAF28105.1; -.
SQ SEQUENCE 197 AA; 21765 MW; BAE0152E18DE7D08 CRC64;

Query Match 99.1%; Score 459; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 3.9e-49;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 RVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLP 171

Qy      62 TPGAFAFHTEFIHVPVGCTCVLPRSV 87
           ||||||||||||||||||||||||||
Db     172 TPGAFAFHTEFIHVPVGCTCVLPRSV 197
```

RESULT 2
Q9NUE6
ID Q9NUE6 PRELIMINARY; PRT; 78 AA.
AC Q9NUE6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   DJ108C2.3 (PUTATIVE NOVEL PROTEIN SIMILAR TO IL17 (INTERLEUKIN 17
DE   (CYTOTOXIC T-LYMPHOCYTE-ASSOCIATED SERINE ESTERASE 8)) (CYTOTOXIC T
DE   LYMPHOCYTE-ASSOCIATED ANTIGEN 8, CTLA8)) (FRAGMENT).
GN   DJ108C2.3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Almeida J.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL034343; CAB75300.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   78 AA;  8704 MW;  66C3C34E7ACC2790 CRC64;


  Query Match             29.4%;  Score 136;  DB 4;  Length 78;
  Best Local Similarity   39.3%;  Pred. No. 1.5e-09;
  Matches   33;  Conservative   14;  Mismatches   23;  Indels   14;  Gaps    3;

Qy       3 VDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRR--PCSRDGSGL 60
           |  | :||| ::  |:|   |||:|: |:|  ::||| : |  ||:||:   ||
Db       1 VTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQGCS------ 53

Qy      61 PTPGAFAFHTEFIHVPVGCTCVLP 84
                :|  | : | |||||| |
Db      54 -----VSFQLEKVLVTVGCTCVTP 72


RESULT   3
O40633
ID   O40633      PRELIMINARY;      PRT;   151 AA.
AC   O40633;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   INTERLEUKIN 17.
OS   Saimiriine herpesvirus 2.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=10381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-488;
RX   MEDLINE=98037620; PubMed=9371569;
RA   Knappe A., Hiller C., Thurau M., Wittmann S., Hofmann H.,
RA   Fleckenstein B., Fickenscher H.;
RT   "The superantigen-homologous viral immediate-early gene ie14/vsag in
RT   herpesvirus saimiri-transformed human T cells.";
RL   J. Virol. 71:9124-9133(1997).
DR   EMBL; Y13183; CAA73627.1; -.
SQ   SEQUENCE   151 AA;  17189 MW;  49B2C9430C46BE32 CRC64;


  Query Match             26.6%;  Score 123;  DB 12;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 1.2e-07;
  Matches   30;  Conservative   12;  Mismatches   25;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:|||  :  |:|   ||::   |     :||| : | :||:|:   ||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHNPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
              :|  | : | |||||| |
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT   4
Q9QXT6
ID   Q9QXT6      PRELIMINARY;      PRT;   180 AA.
AC   Q9QXT6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CYTOKINE-LIKE PROTEIN ZCYTO7.
GN   ZCYTO7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Presnell S., Gilbert T., Whitmore T., Foster D., Hart C., Lehner J.,
RA   Martinez T., Hoffman R., O'Hara P.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF184970; AAF01319.1; -.
SQ   SEQUENCE   180 AA;  20309 MW;  E26F4C72001997C5 CRC64;


  Query Match             23.9%;  Score 110.5;  DB 11;  Length 180;
  Best Local Similarity   33.8%;  Pred. No. 5e-06;
  Matches   27;  Conservative   13;  Mismatches   35;  Indels    5;  Gaps    2;

Qy       3 VDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPT 62
           :: |  | |  |  | ||| ||::  | :|  :: || :  | : :||| | :     |
Db     105 INHDPSRIPADLPEARCLCLGCVNPFTMQEDRSMVSVPVF-SQVPVRRRLCPQP----PR 159

Qy      63 PGAFAFHTEFIHVPVGCTCV 82
           ||          : |||||:
Db     160 PGPCRQRVVMETIAVGCTCI 179


RESULT   5
Q9UHF5
ID   Q9UHF5      PRELIMINARY;      PRT;   180 AA.
AC   Q9UHF5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CYTOKINE-LIKE PROTEIN ZCYTO7 (INTERLEUKIN 20) (INTERLEUKIN 17B).
GN   ZCYTO7 OR IL20.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Presnell S., Gilbert T., Whitmore T., Foster D., Hart C., Lehner J.,
RA   Martinez T., Hoffman R., O'Hara P.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Shi Y., Ullrich S.J., Zhang J., Connolly K., Grzegorzewski K.J.,
RA   Barber M.C., Wang W., Wathen K., Hodge V., Fisher C.L., Olsen H.,
RA   Ruben S.M., Knyazev I., Cho Y.H., Kao V., Wilkinson K.A.,
RA   Carrell J.A., Ebner R.;
RT   "A novel cytokine receptor-ligand pair. Identification, molecular
RT   characterization, and in vivo immunomodulatory activity.";
RL   J. Biol. Chem. 275:19167-19176(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Li H., Chen J., Huang A., Stinson J., Heldens S., Foster J., Dowd P.,
RA   Gurney A.L., Wood W.I.;
RT   "Cloning and Characterization of IL-17B and IL17C, Two New Members of
RT   the IL-17 Cytokine Family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:773-778(2000).
DR   EMBL; AF184969; AAF01318.1; -.
DR   EMBL; AF212311; AAF78775.1; -.
DR   EMBL; AF152098; AAF28104.1; -.
SQ   SEQUENCE   180 AA;  20437 MW;  F1B0BC1446D0B14A CRC64;


  Query Match             21.3%;  Score 98.5;  DB 4;  Length 180;
  Best Local Similarity   32.5%;  Pred. No. 0.00015;
```

```
  Matches   26;  Conservative   12;  Mismatches   37;  Indels    5;  Gaps    2;

Qy      3 VDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPT 62
          :: |  | |  |  | ||| ||::  | :|  :: || :  | : :||| |      |
Db    105 INHDPSRIPVDLPEARCLCLGCVNPFTMQEDRSMVSVPVF-SQVPVRRRLCPPP----PR 159

Qy     63 PGAFAFHTEFIHVPVGCTCV 82
           |          : |||||:
Db    160 TGPCRQRAVMETIAVGCTCI 179


RESULT   6
Q19778
ID   Q19778      PRELIMINARY;      PRT;   148 AA.
AC   Q19778;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   F25D1.3 PROTEIN.
GN   F25D1.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kelly P.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z73973; CAA98268.1; -.
SQ   SEQUENCE   148 AA;  16738 MW;  ADB51F438DFB3940 CRC64;


  Query Match             19.0%;  Score 88;  DB 5;  Length 148;
  Best Local Similarity   32.6%;  Pred. No. 0.0025;
  Matches   28;  Conservative   15;  Mismatches   31;  Indels   12;  Gaps    4;

Qy      1 ARVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGL 60
          :||:  | | |: :| : |||      |  | :     : ::: : :|||  | |  :||
Db     59 SRVNYQESREPKLIAESVCLC------RKSRGSTGAFCMPIVRKVPILRRVSCDR-STGL 111

Qy     61 PTPGAFAFHTEFIHVPVGCTCVLPRS 86
              :   || |   |||  ||||:
Db    112 ---WNYVRSTELI--TVGCHSVLPRT 132


RESULT   7
Q9TUP5
ID   Q9TUP5      PRELIMINARY;      PRT;    39 AA.
AC   Q9TUP5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   INTERLEUKIN 17 (FRAGMENT).
GN   IL-17.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Laurent F., Lacroix S.M., Mancassola R., Naciri M.;
RT   "Partial sequence of bovine IL-17.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF166487; AAD46376.1; -.
FT   NON_TER       1      1
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4394 MW;  4A390A22973482A3 CRC64;


  Query Match             15.0%;  Score 69.5;  DB 6;  Length 39;
  Best Local Similarity   42.5%;  Pred. No. 0.13;
  Matches   17;  Conservative    7;  Mismatches   15;  Indels    1;  Gaps    1;

Qy      9 RYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVL 48
          |||  :  |:|   |||:|  |:    :||| : | :|||
Db      1 RYPSVIWEAKCSHSGCINAE-GKVDHHMNSVTIQQEILVL 39


RESULT   8
Q9TUP4
ID   Q9TUP4      PRELIMINARY;      PRT;    39 AA.
AC   Q9TUP4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   INTERLEUKIN 17 (FRAGMENT).
GN   IL-17.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Laurent F., Lacroix S.M., Mancassola R., Naciri M.;
RT   "Partial sequence of ovine IL-17.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF166488; AAD46377.1; -.
FT   NON_TER       1      1
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4394 MW;  4A390A22973482A3 CRC64;


  Query Match             15.0%;  Score 69.5;  DB 6;  Length 39;
  Best Local Similarity   42.5%;  Pred. No. 0.13;
  Matches   17;  Conservative    7;  Mismatches   15;  Indels    1;  Gaps    1;

Qy      9 RYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVL 48
          |||  :  |:|   |||:|  |:    :||| : | :|||
Db      1 RYPSVIWEAKCSHSGCINAE-GKVDHHMNSVTIQQEILVL 39


RESULT   9
Q9TUP3
ID   Q9TUP3      PRELIMINARY;      PRT;    39 AA.
AC   Q9TUP3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   INTERLEUKIN 17 (FRAGMENT).
GN   IL-17.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Laurent F., Lacroix S.M., Mancassola R., Naciri M.;
RT   "Partial sequence of porcine IL-17.";
```

```
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF166489; AAD46378.1; -.
FT   NON_TER       1      1
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4394 MW;  4A390A22973482A3 CRC64;


  Query Match             15.0%;  Score 69.5;  DB 6;  Length 39;
  Best Local Similarity   42.5%;  Pred. No. 0.13;
  Matches   17;  Conservative    7;  Mismatches   15;  Indels    1;  Gaps    1;

Qy      9 RYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVL 48
          |||   :  |:|   |||:|  |     :||| : | :|||
Db      1 RYPSVIWEAKCSHSGCINAE-GRVDHHMNSVTIQQEILVL 39


RESULT  10
Q9S2C5
ID   Q9S2C5      PRELIMINARY;      PRT;   146 AA.
AC   Q9S2C5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL 15.0 KDA PROTEIN.
GN   SCF85.16.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Seeger K.J., Harris D.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
DR   EMBL; AL110470; CAB54173.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   146 AA;  14974 MW;  4350BBBEF04315B6 CRC64;


  Query Match             14.5%;  Score 67;  DB 2;  Length 146;
  Best Local Similarity   43.1%;  Pred. No. 0.99;
  Matches   22;  Conservative    2;  Mismatches   15;  Indels   12;  Gaps    3;

Qy     25 IDARTGRETAALNSVRLLQSLLVLRRRPCSRDGS-GLP---TPGAFAFHTE 71
          : ||| :   ||  ||         ||| ||||| | |    || |   ||
Db    102 LQARTPQTAQALVGVR--------ARRPLSRDGSAGAPRGRRPGLFLTRTE 144


RESULT  11
O13035
ID   O13035      PRELIMINARY;      PRT;   518 AA.
AC   O13035;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PROACTIVATOR POLYPEPTIDE PRECURSOR [CONTAINS: SAPOSIN A (PROTEIN A);
DE   SAPOSIN B (SPHINGOLIPID ACTIVATOR PROTEIN 1) (SAP-1) (DISPERSIN)
DE   (SULFATIDE/GM1 ACTIVATOR); SAPOSIN C (CO-BETA-GLUCOSIDASE) (A1
DE   ACTIVATOR) (GLUCOSYLCERAMIDASE ACTIVATOR) (SPHINGOLIPID ACTIVATOR
DE   PROTEIN 2) (SAP-2); SAPOSIN D (PROTEIN C) (COMPONENT C)].
GN   PSAP.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.
RC   TISSUE=BRAIN, AND LIVER;
RX   MEDLINE=98129745; PubMed=9461526;
RA   Azuma N., Seo H.-C., Lie O., Fu Q., Gould R.M., Hiraiwa M., Burt D.W.,
RA   Paton I.R., Morrice D.R., O Brien J.S., Kishimoto Y.;
RT   "Cloning, expression and map assignment of chicken prosaposin.";
RL   Biochem. J. 330:321-327(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Altman N., Horowitz M.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE LYSOSOMAL DEGRADATION OF SPHINGOLIPIDS TAKES PLACE
CC       BY THE SEQUENTIAL ACTION OF SPECIFIC HYDROLASES. SOME OF THESE
CC       ENZYMES REQUIRE SPECIFIC LOW-MOLECULAR MASS, NON-ENZYMIC PROTEINS:
CC       THE SPHINGOLIPIDS ACTIVATOR PROTEINS (COPROTEINS) (BY SIMILARITY).
CC   -!- FUNCTION: SAPOSIN A AND SAPOSIN C STIMULATE THE HYDROLYSIS OF
CC       GLUCOSYLCERAMIDE BY BETA-GLUCOSYLCERAMIDASE (EC 3.2.1.45) AND
CC       GALACTOSYLCERAMIDE BY BETA-GALACTOSYLCERAMIDASE (EC 3.2.1.46).
CC       SAPOSIN-C APPARENTLY ACTS BY COMBINING WITH THE ENZYME AND ACIDIC
CC       LIPID TO FORM AN ACTIVATED COMPLEX, RATHER THAN BY SOLUBILIZING
CC       THE SUBSTRATE (BY SIMILARITY).
CC   -!- FUNCTION: SAPOSIN B STIMULATES THE HYDROLYSIS OF GALACTO-
CC       CEREBROSIDE SULFATE BY ARYLSULFATASE A (EC 3.1.6.8), GM1
CC       GANGLIOSIDES BY BETA-GALACTOSIDASE (EC 3.2.1.23) AND
CC       GLOBOTRIAOSYLCERAMIDE BY ALPHA-GALACTOSIDASE A (EC 3.2.1.22).
CC       SAPOSIN-B FORMS A SOLUBILIZING COMPLEX WITH THE SUBSTRATES OF THE
CC       SPHINGOLIPID HYDROLASES (BY SIMILARITY).
CC   -!- FUNCTION: SAPOSIN D IS A SPECIFIC SPHINGOMYELIN PHOSPHODIESTERASE
CC       ACTIVATOR (EC 3.1.4.12) (BY SIMILARITY).
CC   -!- SUBUNIT: SAPOSIN B IS A HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL (BY SIMILARITY).
CC   -!- PTM: THIS PRECURSOR IS PROTEOLYTICALLY PROCESSED TO 4 SMALL
CC       PEPTIDES, WHICH ARE SIMILAR TO EACH OTHER AND ARE SPHINGOLIPID
CC       HYDROLASE ACTIVATOR PROTEINS (BY SIMILARITY).
CC   -!- PTM: THERE ARE THREE DISULFIDE BRIDGES IN EACH OF THE SAPOSINS.
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS FOUR SAPOSIN/SP-B B-TYPE DOMAINS AND TWO
CC       A-TYPE DOMAINS. THE B-TYPE DOMAINS CORRESPOND TO THE MATURE
CC       SAPOSINS.
DR   EMBL; AB003471; BAA19914.1; -.
DR   EMBL; AF108656; AAF05899.1; -.
DR   INTERPRO; IPR000004; -.
DR   INTERPRO; IPR003119; -.
DR   INTERPRO; IPR003259; -.
DR   PFAM; PF02199; SAPA; 2.
DR   PRODOM; PD001732; -; 3.
DR   PRODOM; PD012321; -; 1.
KW   Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;
KW   GM2-gangliosidosis.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        61    143       SAPOSIN A.
FT   CHAIN       194    276       SAPOSIN B.
FT   CHAIN       307    387       SAPOSIN C.
FT   CHAIN       399    480       SAPOSIN D.
FT   DOMAIN       22     55       SAPOSIN/SP-B A-TYPE DOMAIN.
FT   DOMAIN      485    518       SAPOSIN/SP-B A-TYPE DOMAIN.
FT   CARBOHYD     81     81       POTENTIAL.
FT   CARBOHYD    214    214       POTENTIAL.
FT   CARBOHYD    328    328       POTENTIAL.
FT   CARBOHYD    420    420       POTENTIAL.
FT   CONFLICT     94     94       R -> T (IN REF. 2).
FT   CONFLICT    486    486       E -> D (IN REF. 2).
SQ   SEQUENCE   518 AA;  57601 MW;  B803000E891C3963 CRC64;


  Query Match             14.0%;  Score 65;  DB 13;  Length 518;
```

```
  Best Local Similarity   27.8%;  Pred. No. 6.3;
  Matches   15;  Conservative   11;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       7 EDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGL 60
           :|:  ||    | :|: ||   |  : |   :   ::||:   :  | | | |:
Db     183 QDKPKQKSKATEDVCQDCIRLVTDVQEAVRTNATFVKSLVAHAKEECDRLGPGM 236


RESULT  12
O24596
ID   O24596      PRELIMINARY;      PRT;   563 AA.
AC   O24596;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PECTIN METHYLESTERASE-LIKE PROTEIN.
GN   C5.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A188; TISSUE=POLLEN;
RX   MEDLINE=98281586; PubMed=9620276;
RA   Wakeley P.R., Rogers H.J., Rozycka M., Greenland A.J., Hussey P.J.;
RT   "A maize pectin methylesterase-like gene, ZmC5, specifically expressed
RT   in pollen.";
RL   Plant Mol. Biol. 37:187-192(1998).
DR   EMBL; Y13285; CAA73733.1; -.
DR   MENDEL; 27043; Zeama;1273;27043.
DR   INTERPRO; IPR000070; -.
DR   PFAM; PF01095; Pectinesterase; 1.
DR   PROSITE; PS00503; PECTINESTERASE_2; 1.
SQ   SEQUENCE   563 AA;  61711 MW;  1F3BD3457B5BBE53 CRC64;


  Query Match             13.7%;  Score 63.5;  DB 10;  Length 563;
  Best Local Similarity   27.3%;  Pred. No. 10;
  Matches   18;  Conservative   10;  Mismatches   29;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDR---YPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGS 58
           | |  :|     |::  |  |:  | ||   |      ||  : |: |:: |||     :
Db     373 RFDAFQDTLYVQPRRQFFRNCVVSGTIDFIFG------NSAAVFQNCLIITRRPMDNQQN 426

Qy      59 GLPTPG 64
           :   |
Db     427 SVTAHG 432


RESULT  13
Q9U195
ID   Q9U195      PRELIMINARY;      PRT;   745 AA.
AC   Q9U195;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL 81.3 KDA PROTEIN.
GN   L2903.03.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Wedler H., Duesterhoeft A., Ivens A.C., Lawson D., Quail M.,
RA   Rajandream M.A., Barrell B.G.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA   Smith D.F.;
RT   "A physical map of the Leishmania major Friedlin genome.";
RL   Genome Res. 8:135-145(1998).
DR   EMBL; AL117319; CAB55517.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   745 AA;  81343 MW;  1C52ABD8DB24D935 CRC64;


  Query Match             13.5%;  Score 62.5;  DB 5;  Length 745;
  Best Local Similarity   33.3%;  Pred. No. 18;
  Matches   18;  Conservative    6;  Mismatches   19;  Indels   11;  Gaps    2;

Qy      15 AFAECLCR---------GCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSG 59
           |  | | |         | : | |  :  | |   :| :    ::|||| ||||
Db     172 AVVEALVRLVPDSLFTTGAVPASTATDGGASNIFAVLSA--YAQQRPCSEDGSG 223


RESULT  14
O75097
ID   O75097      PRELIMINARY;      PRT;  1737 AA.
AC   O75097;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   MEGF8 (FRAGMENT).
GN   MEGF8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE=98360089; PubMed=9693030;
RA   Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT   "Identification of high-molecular-weight proteins with multiple EGF-
RT   like motifs by motif-trap screening.";
RL   Genomics 51:27-34(1998).
DR   EMBL; AB011541; BAA32469.1; -.
DR   HSSP; P02468; 1TLE.
DR   INTERPRO; IPR000152; -.
DR   INTERPRO; IPR000561; -.
DR   INTERPRO; IPR000859; -.
DR   INTERPRO; IPR001881; -.
DR   INTERPRO; IPR002049; -.
DR   PFAM; PF00053; laminin_EGF; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_4.
DR   PROSITE; PS01180; CUB; 1.
DR   PROSITE; PS01186; EGF_2; 4.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; UNKNOWN_4.
FT   NON_TER       1      1
SQ   SEQUENCE   1737 AA;  184680 MW;  FD3152E415193159 CRC64;


  Query Match             13.4%;  Score 62;  DB 4;  Length 1737;
  Best Local Similarity   27.8%;  Pred. No. 50;
  Matches   20;  Conservative    7;  Mismatches   33;  Indels   12;  Gaps    2;

Qy      19 CLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPC------SRDGSGLPTPGAFAFHTEF 72
           |:  ||     | |: :|   :  |  | |||  |       :|| |    |  :   :
Db     994 CMAGGCGRLLRGPESCSLGCAQATQCALCLRRPHCGWCAWGGQDGGGRCMEGGLSGPRD- 1052

Qy      73 IHVPVGCTCVLP 84
                | ||  |
Db    1053 -----GLTCGRP 1059


RESULT  15
Q9QYP0
ID   Q9QYP0      PRELIMINARY;      PRT;   874 AA.
AC   Q9QYP0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
```

```
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   MEGF8 (FRAGMENT).
GN   MEGF8.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RA   Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT   "Identification of high-molecular-weight proteins with multiple EGF-
RT   like motifs by motif-trap screening.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB011534; BAA88689.1; -.
DR   HSSP; P35555; 1EMN.
DR   INTERPRO; IPR000152; -.
DR   INTERPRO; IPR000561; -.
DR   INTERPRO; IPR001881; -.
DR   INTERPRO; IPR002049; -.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; UNKNOWN_2.
FT   NON_TER       1       1
SQ   SEQUENCE   874 AA;  93764 MW;  CB63BDF183D870AB CRC64;


  Query Match             13.2%;  Score 61;  DB 11;  Length 874;
  Best Local Similarity   27.8%;  Pred. No. 33;
  Matches   20;  Conservative    7;  Mismatches   33;  Indels   12;  Gaps    2;

Qy      19 CLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPC------SRDGSGLPTPGAFAFHTEF 72
           |:  ||     | |: :|   :  |  | |||  |       :|| |    |  :   :
Db     131 CMAGGCGRLLRGPESCSLGCAQATQCALCLRRPHCGWCAWGGQDGGGHCMEGGLSGPRD- 189

Qy      73 IHVPVGCTCVLP 84
              | ||  |
Db     190 -----GLTCGRP 196


Search completed: February 16, 2001, 12:52:39
Job time: 194 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:49:50 ; Search time 14.73 Seconds
(without alignments)
190.739 Million cell updates/sec

Title: US-09-320-713-2
Perfect score: 463
Sequence: 1 ARVDTDEDRYPQKLAFAECL.....FHTEFIHVPVGCTCVLPRSV 87

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	125	27.0	155	1	IL17_HUMAN	Q16552 homo sapien
2	124	26.8	151	1	VG13_HSVSA	P24916 herpesvirus
3	119	25.7	150	1	IL17_RAT	Q61453 rattus norv
4	119	25.7	158	1	IL17_MOUSE	Q62386 mus musculu
5	63	13.6	362	1	A2HS_PIG	P29700 sus scrofa
6	62.5	13.5	762	1	P115_CHICK	Q98917 gallus gall
7	60.5	13.1	306	1	YE16_YEAST	P40101 saccharomyc
8	60.5	13.1	909	1	CT1A_FUSSO	P52958 fusarium so
9	60	13.0	344	1	YD33_MYCTU	Q10644 mycobacteri
10	60	13.0	454	1	DAT_HAEIN	P44951 haemophilus
11	58.5	12.6	640	1	CLAT_PIG	P13222 sus scrofa
12	57.5	12.4	452	1	CLPX_HELPJ	Q9zj18 helicobacte
13	57	12.3	303	1	YJ85_MYCTU	Q10872 mycobacteri
14	56.5	12.2	455	1	AROA_AGRSP	Q9r4e4 agrobacteri
15	56	12.1	398	1	CAR4_RHINI	Q03700 rhizopus ni
16	56	12.1	798	1	YLFN_CAEEL	Q20296 caenorhabdi
17	56	12.1	1964	1	NTC4_MOUSE	P31695 mus musculu
18	55.5	12.0	298	1	MIAA_AGRTU	P38436 agrobacteri
19	55.5	12.0	328	1	PIT_BACSU	O34436 bacillus su
20	55.5	12.0	363	1	D1DR_CARAU	P35406 carassius a
21	55.5	12.0	445	1	DAT_ACIBA	P56744 acinetobact
22	55.5	12.0	547	1	PPB_BOMMO	P29523 bombyx mori
23	55	11.9	350	1	REDD_STRCO	P16922 streptomyce
24	55	11.9	524	1	SAP_HUMAN	P07602 h proactiva
25	55	11.9	635	1	GIDA_SYNY3	Q55694 synechocyst
26	54.5	11.8	470	1	RHSA_RHIME	Q9z3r2 rhizobium m
27	54.5	11.8	491	1	ACHE_BOVIN	P02715 bos taurus
28	54.5	11.8	493	1	ACHE_HUMAN	Q04844 homo sapien
29	54.5	11.8	550	1	IDS_HUMAN	P22304 homo sapien
30	54.5	11.8	782	1	ZFY1_MOUSE	P10925 mus musculu
31	54.5	11.8	783	1	ZFY2_MOUSE	P20662 mus musculu
32	54	11.7	305	1	Y441_TREPA	O83455 treponema p
33	53.5	11.6	446	1	CLPX_HELPY	O25926 helicobacte
34	53.5	11.6	499	1	KPY1_TRYBB	P30615 trypanosoma
35	53.5	11.6	499	1	KPY2_TRYBB	P30616 trypanosoma
36	53.5	11.6	1587	1	TOP2_PENCH	Q9y8g8 penicillium
37	53.5	11.6	1687	1	Z142_HUMAN	P52746 homo sapien
38	53	11.4	412	1	NH64_CAEEL	O44960 caenorhabdi
39	53	11.4	456	1	YRDX_RHOSH	Q01856 rhodobacter
40	53	11.4	3051	1	YNX3_CAEEL	P34576 caenorhabdi
41	53	11.4	5032	1	RYNR_HUMAN	P21817 homo sapien
42	52.5	11.3	278	1	MA32_MOUSE	O35658 mus musculu
43	52.5	11.3	495	1	CD5_HUMAN	P06127 homo sapien
44	52.5	11.3	542	1	HM18_CAEEL	P41934 caenorhabdi
45	52.5	11.3	785	1	VD05_VACCC	P21010 vaccinia vi

ALIGNMENTS

RESULT 1
IL17_HUMAN
ID IL17_HUMAN STANDARD; PRT; 155 AA.
AC Q16552;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED
DE ANTIGEN 8) (CTLA-8).
GN IL17 OR CTLA8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96281911; PubMed=8676080;
RA Fossiez F., Djossou O., Chomarat P., Flores-Romo L., Ait-Yahia S.,
RA Maat C., Pin J.-J., Garrone P., Garcia E., Saeland S., Blanchard D.,
RA Gaillard C., Das Mahapatra B., Rouvier E., Golstein P., Banchereau J.,
RA Lebecque S.;
RT "T cell interleukin-17 induces stromal cells to produce
RT proinflammatory and hematopoietic cytokines.";
RL J. Exp. Med. 183:2593-2603(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=T-CELL;
RX MEDLINE=96094436; PubMed=7499828;
RA Yao Z., Painter S.L., Fanslow W.C., Ulrich D., Macduff B.M.,
RA Spriggs M.K., Armitage R.J.;
RT "Human IL-17: a novel cytokine derived from T cells.";
RL J. Immunol. 155:5483-5486(1995).
CC -!- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND
CC HEMATOPOIETIC CYTOKINES. ENHANCES THE SURFACE EXPRESSION OF THE
CC INTRACELLULAR ADHESION MOLECULE-1 (ICAM-1) IN FIBROBLASTS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: RESTRICTED TO ACTIVATED MEMORY T-CELLS.
CC -!- PTM: FOUND BOTH IN GLYCOSYLATED AND NONGLYCOSYLATED FORMS.
CC -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO
CC HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.
CC ---

CC ---
DR EMBL; Z58820; CAA91233.1; -.
DR EMBL; U32659; AAC50341.1; -.
DR MIM; 603149; -.
KW Cytokine; Glycoprotein; T-cell; Antigen; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 155 INTERLEUKIN-17.
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
SQ   SEQUENCE   155 AA;  17504 MW;  2BCAE9CB2F4886D1 CRC64;

  Query Match             27.0%;  Score 125;  DB 1;  Length 155;
  Best Local Similarity   40.7%;  Pred. No. 2.1e-08;
  Matches   33;  Conservative    8;  Mismatches   26;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRP--CSRDGSGLPTP 63
           | :|||  :  |:|   |||:|  |     :||| : | :||||| |  |
Db      81 DPERYPSVIWEAKCRHLGCINA-DGNVDYHMNSVPIQQEILVLRREPPHCPN-------- 131

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | | | |||||| |
Db     132 ---SFRLEKILVSVGCTCVTP 149


RESULT   2
VG13_HSVSA
ID   VG13_HSVSA     STANDARD;      PRT;   151 AA.
AC   P24916;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   IMMEDIATE EARLY GENE 13 PROTEIN PRECURSOR.
GN   13 OR KCLF2.
OS   Herpesvirus saimiri (strain 11).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90163221; PubMed=2154888;
RA   Albrecht J.-C., Fleckenstein B.;
RT   "Structural organization of the conserved gene block of Herpesvirus
RT   saimiri coding for DNA polymerase, glycoprotein B, and major DNA
RT   binding protein.";
RL   Virology 174:533-542(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92333688; PubMed=1321287;
RA   Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA   Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA   Honess R.W.;
RT   "Primary structure of the herpesvirus saimiri genome.";
RL   J. Virol. 66:5047-5058(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91021021; PubMed=1699352;
RA   Nicholas J., Smith E.P., Coles L., Honess R.;
RT   "Gene expression in cells infected with gammaherpesvirus saimiri:
RT   properties of transcripts from two immediate-early genes.";
RL   Virology 179:189-200(1990).
CC   -!- SIMILARITY: STRONG, TO MAMMALIAN INTERLEUKIN-17 (CTLA-8).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X64346; CAA45636.1; -.
DR   EMBL; M31122; AAA46169.1; -.
DR   EMBL; M60286; AAA46156.1; -.
DR   PIR; D36807; D36807.
DR   PIR; B45351; B45351.
KW   Cytokine; Early protein; Signal.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    151       IMMEDIATE EARLY GENE 13 PROTEIN.
FT   CARBOHYD    36     36       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    53     53       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   151 AA;  17180 MW;  53BEDDE4206C6432 CRC64;

  Query Match             26.8%;  Score 124;  DB 1;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 2.7e-08;
  Matches   30;  Conservative   13;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:||||  :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||||| |
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT   3
IL17_RAT
ID   IL17_RAT       STANDARD;      PRT;   150 AA.
AC   Q61453;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED
DE   ANTIGEN 8) (CTLA-8).
GN   IL17 OR CTLA8.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93294300; PubMed=8390535;
RA   Rouvier E., Luciani M.F., Mattei M.-G., Denizot F., Golstein P.;
RT   "CTLA-8, cloned from an activated T cell, bearing AU-rich messenger
RT   RNA instability sequences, and homologous to a herpesvirus saimiri
RT   gene.";
RL   J. Immunol. 150:5445-5456(1993).
RN   [2]
RP   ORGANISM IDENTIFICATION.
RX   MEDLINE=96194901; PubMed=8654948;
RA   Yao Z., Timour M., Painter S., Fanslow W., Spriggs M.K.;
RT   "Complete nucleotide sequence of the mouse CTLA8 gene.";
RL   Gene 168:223-225(1996).
RN   [3]
RP   SEQUENCE FROM N.A., AND ORGANISM IDENTIFICATION.
RX   MEDLINE=97031826; PubMed=8877732;
RA   Kennedy J., Rossi D.L., Zurawski S.M., Vega F. Jr., Kastelein R.A.,
RA   Wagner J.L., Hannum C.H., Zlotnik A.;
RT   "Mouse IL-17: a cytokine preferentially expressed by alpha beta TCR +
RT   CD4-CD8-T cells.";
RL   J. Interferon Cytokine Res. 16:611-617(1996).
CC   -!- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND
CC       HEMATOPOIETIC CYTOKINES (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO
CC       HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.
CC   -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE FROM MOUSE BUT, ON
CC       THE BASIS OF SUBSEQUENT WORK (REF.2 AND REF.3), HAS BEEN SHOWN TO
CC       BE OF RAT ORIGIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L13839; AAA37490.1; -.
KW   Cytokine; Glycoprotein; T-cell; Antigen; Signal.
FT   SIGNAL       1     17       POTENTIAL.
FT   CHAIN       18    150       INTERLEUKIN-17.
FT   CARBOHYD    63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CONFLICT     46     46       I -> L (IN REF. 3).
SQ   SEQUENCE   150 AA;  16876 MW;  EF13F33EDF9D689F CRC64;


  Query Match             25.7%;  Score 119;  DB 1;  Length 150;
  Best Local Similarity   38.0%;  Pred. No. 1.1e-07;
  Matches   30;  Conservative   11;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRP--CSRDGSGLPTP 63
           | ||||  :  |:|  : |::|  |:    :||| : | :|||:| |  |
Db      76 DPDRYPSVIWEAQCRHQRCVNAE-GKLDHHMNSVLIQQEILVLKREPEKC---------- 124

Qy      64 GAFAPHTEFIHVPVGCTCV 82
             | |  | : | ||||||
Db     125 -PFTFRVEKMLVGVGCTCV 142


RESULT   4
IL17_MOUSE
ID   IL17_MOUSE     STANDARD;      PRT;   158 AA.
AC   Q62386; Q60971;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED
DE   ANTIGEN 8) (CTLA-8).
GN   IL17 OR CTLA8.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=THYMOCYTES;
RX   MEDLINE=97031826; PubMed=8877732;
RA   Kennedy J., Rossi D.L., Zurawski S.M., Vega F. Jr., Kastelein R.A.,
RA   Wagner J.L., Hannum C.H., Zlotnik A.;
RT   "Mouse IL-17: a cytokine preferentially expressed by alpha beta TCR +
RT   CD4-CD8-T cells.";
RL   J. Interferon Cytokine Res. 16:611-617(1996).
RN   [2]
RP   SEQUENCE OF 12-158 FROM N.A.
RC   STRAIN=129/SV; TISSUE=T-CELL;
RX   MEDLINE=96194901; PubMed=8654948;
RA   Yao Z., Timour M., Painter S., Fanslow W., Spriggs M.K.;
RT   "Complete nucleotide sequence of the mouse CTLA8 gene.";
RL   Gene 168:223-225(1996).
CC   -!- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND
CC       HEMATOPOIETIC CYTOKINES (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO A SUBSET OF ACTIVATED T-CELLS.
CC   -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO
CC       HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U43088; AAB05222.1; -.
DR   EMBL; U35108; AAA93253.1; -.
DR   MGD; MGI:107364; IL17.
KW   Cytokine; Glycoprotein; T-cell; Antigen; Signal.
FT   SIGNAL       1     25       POTENTIAL.
FT   CHAIN       26    158       INTERLEUKIN-17.
FT   CARBOHYD    71     71       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   158 AA;  17490 MW;  3505C143435F4653 CRC64;


  Query Match             25.7%;  Score 119;  DB 1;  Length 158;
  Best Local Similarity   39.0%;  Pred. No. 1.2e-07;
  Matches   30;  Conservative   11;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGA 65
           | ||||  :  |:|  : |::|  |:    :||| : | :|||:| | |
Db      84 DPDRYPSVIWEAQCRHQRCVNAE-GKLDHHMNSVLIQQEILVLKREPES---------CP 133

Qy      66 FAPHTEFIHVPVGCTCV 82
             | |  | : | ||||||
Db     134 FTFRVEKMLVGVGCTCV 150


RESULT   5
A2HS_PIG
ID   A2HS_PIG       STANDARD;      PRT;   362 AA.
AC   P29700;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   ALPHA-2-HS-GLYCOPROTEIN PRECURSOR (FETUIN) (FRAGMENT).
GN   AHSG.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=LIVER;
RX   MEDLINE=92209519; PubMed=1372866;
RA   Brown W.M., Christie D.L., Saunders N.R., Nawratil P.,
RA   Dziegielewska K.D., Mueller-Esterl W.;
RT   "The nucleotide and deduced amino acid structures of sheep and pig
RT   fetuin. Common structural features of the mammalian fetuin family.";
RL   Eur. J. Biochem. 205:321-331(1992).
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE FETUIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 CYSTATIN-LIKE REPEATS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X56021; CAA39498.1; -.
DR   PIR; S22142; S22142.
DR   PIR; S22395; S22395.
DR   INTERPRO; IPR000010; -.
DR   INTERPRO; IPR001363; -.
DR   PFAM; PF00031; cystatin; 2.
DR   PROSITE; PS01254; FETUIN_1; 1.
DR   PROSITE; PS01255; FETUIN_2; 1.
KW   Glycoprotein; Signal; Repeat.
FT   NON_TER      1      1
FT   SIGNAL      <1     15
FT   CHAIN       16    362       ALPHA-2-HS-GLYCOPROTEIN.
FT   REPEAT      24    141       CYSTATIN-LIKE.
FT   REPEAT     142    258       CYSTATIN-LIKE.
FT   DISULFID    29    353       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    86     97       BY SIMILARITY.
FT   DISULFID   111    129       BY SIMILARITY.
FT   DISULFID   143    146       BY SIMILARITY.
FT   DISULFID   205    216       BY SIMILARITY.
FT   DISULFID   227    244       BY SIMILARITY.
FT   CARBOHYD    96     96       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   153    153       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   173    173       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   362 AA;  38424 MW;  532648EE434B5686 CRC64;


  Query Match             13.6%;  Score 63;  DB 1;  Length 362;
  Best Local Similarity   25.4%;  Pred. No. 2.3;
```

```
  Matches   17;  Conservative    8;  Mismatches   42;  Indels    0;  Gaps    0;

Qy      21 CRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGAFAFHTEFIHVPVGCT 80
           |:| : |:   |  |:         :||: :|   |    |   | |  :    ||
Db     227 CKGTVTAKVNEEDVAVTCTVFQTQPVVLQPQPAGADAGATPVVDAAATASPLADVPAASL 286

Qy      81 CVLPRSV 87
           | |  |
Db     287 VVGPMVV 293


RESULT   6
P115_CHICK
ID   P115_CHICK     STANDARD;      PRT;   762 AA.
AC   Q98917;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MELANOSOMAL MATRIX PROTEIN 115 KDA PROTEIN PRECURSOR.
GN   MMP115.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN; TISSUE=RETINAL PIGMENT EPITHELIUM;
RX   MEDLINE=92020667; PubMed=1924173;
RA   Mochii M., Agata K., Eguchi G.;
RT   "Complete sequence and expression of a cDNA encoding a chicken
RT   115-kDa melanosomal matrix protein.";
RL   Pigment Cell Res. 4:41-47(1991).
RN   [2]
RP   CHARACTERIZATION.
RC   STRAIN=WHITE LEGHORN; TISSUE=RETINAL PIGMENT EPITHELIUM;
RX   MEDLINE=88311098; PubMed=3409326;
RA   Mochii M., Agata K., Kobayashi H., Yamamoto T.S., Eguchi G.;
RT   "Expression of gene coding for a melanosomal matrix protein
RT   transcriptionally regulated in the transdifferentiation of chick
RT   embryo pigmented epithelial cells.";
RL   Cell Differ. 24:67-74(1988).
CC   -!- FUNCTION: MIGHT BE REQUIRED FOR POLYMERIZATION OF MELANIN ONTO THE
CC       CORE STRUCTURE OF MELANOSOMES WITH ENZYMIC FUNCTION OF TYROSINASE.
CC   -!- SUBCELLULAR LOCATION: ON THE FIBROUS MATRIX STRUCTURE OF THE
CC       PREMELANOSOME.
CC   -!- TISSUE SPECIFICITY: SPECIFIC TO PIGMENTED EPITHELIAL CELLS AND
CC       MELANOCYTES. NOT EXPRESSED IN LENS, NEURAL RETINA, BRAIN, HEART,
CC       GIZZARD OR LIVER.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSEDED DURING THE REDIFFERENTIATION OF
CC       PIGMENTED EPITHELIAL CELLS (PEC).
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 PKD DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D88348; BAA13589.1; -.
DR   INTERPRO; IPR000601; -.
DR   PFAM; PF00801; PKD; 1.
DR   PROSITE; PS50093; PKD; 1.
KW   Signal; Glycoprotein; Repeat.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    762       MELANOSOMAL MATRIX PROTEIN 115 KDA
FT                                PROTEIN.
FT   DOMAIN      223    323       PKD.
FT   DOMAIN      441    532       4 X 20-24 AA APPROXIMATE TANDEM REPEATS.
FT   REPEAT      441    464       1.
FT   REPEAT      465    488       2.
FT   REPEAT      489    508       3.
FT   REPEAT      509    532       4.
FT   CARBOHYD    111    111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    346    346       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    651    651       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    659    659       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   762 AA;  77356 MW;  172C8DDB4FDCE7C6 CRC64;


  Query Match             13.5%;  Score 62.5;  DB 1;  Length 762;
  Best Local Similarity   26.2%;  Pred. No. 5.9;
  Matches   17;  Conservative   13;  Mismatches   24;  Indels   11;  Gaps    1;

Qy      23 GCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGAFAFHTEFIHVPVGCTCV 82
           ||:  | |  :  || |: ::|: :::  | : :|||            : : | |
Db     563 GCVLYRYGTFSTELNIVQGIESVAIVQVVPAAPEGSG----------NSVELTVTCEGS 611

Qy      83 LPRSV 87
           ||  |
Db     612 LPEEV 616


RESULT   7
YE16_YEAST
ID   YE16_YEAST     STANDARD;      PRT;   306 AA.
AC   P40101;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   HYPOTHETICAL 35.9 KDA PROTEIN IN ISC10 3'REGION.
GN   YER186C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC   Saccharomycetaceae; Saccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA   Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA   Chung E., Duncan M., Guzman E., Hartzell G., Hunicke-Smith S.,
RA   Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA   Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA   Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA   Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18922; AAB64713.1; -.
DR   SGD; S0000988; YER186C.
KW   Hypothetical protein.
SQ   SEQUENCE   306 AA;  35911 MW;  9A024CD8116E4E5A CRC64;


  Query Match             13.1%;  Score 60.5;  DB 1;  Length 306;
  Best Local Similarity   31.0%;  Pred. No. 3.9;
  Matches   22;  Conservative    9;  Mismatches   31;  Indels    9;  Gaps    4;

Qy       6 DEDR--YPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRP---CSRDGSGL 60
             :|||   | |  |       | :  |  ::     | | : :|   |: |    : || ||
Db      22 NEDRKDVPDKFKFDTVTKTGMVKLRVFKDDLIFKSQRSI-NLFASRKHPFKSFTADGEGL 80

Qy      61 PTPGAFAFHTE 71
           |    ||| |:
Db      81 P---LFAFRTK 88
```

```
RESULT   8
CT1A_FUSSO
ID   CT1A_FUSSO     STANDARD;      PRT;   909 AA.
AC   P52958;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CUTINASE TRANSCRIPTION FACTOR 1 ALPHA.
GN   CTF1-ALPHA.
OS   Fusarium solani (subsp. pisi) (Nectria haematococca).
OC   Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Fusarium.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Li D., Kolattukudy P.E.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC       CLUSTER DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U51671; AAA96824.1; -.
DR   INTERPRO; IPR001138; -.
DR   PFAM; PF00172; Zn_clus; 1.
DR   PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR   PROSITE; PS50048; ZN2_CY6_FUNGAL_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein; Zinc;
KW   Metal-binding.
FT   DNA_BIND     61     90       ZN(2)-CYS(6), FUNGAL-TYPE.
SQ   SEQUENCE   909 AA;  101172 MW;  66AC2EC18F5C96A9 CRC64;

  Query Match             13.1%;  Score 60.5;  DB 1;  Length 909;
  Best Local Similarity   28.9%;  Pred. No. 13;
  Matches   24;  Conservative    8;  Mismatches   26;  Indels   25;  Gaps    3;

Qy      12 QKLAF---AECLCRGCIDA-----------------RTGRETAALNSVRLLQSLLVLRR 50
           |: ||      ||   |||                 |  |:     |: ||||:|:
Db     240 QRGAFLLPPRALCDELIDAYFSWVHPIVPVINRTRFMRQYRDPKNPPSLLLLQSVLLAGT 299

Qy      51 RPCSR----DGSGLPTPGAFAFH 69
           | |:     | :|  || |  |:
Db     300 RACNNAQLMDANGSTTPAALTFY 322


RESULT   9
YD33_MYCTU
ID   YD33_MYCTU     STANDARD;      PRT;   344 AA.
AC   Q10644;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL 33.9 KDA PROTEIN RV1333.
GN   RV1333 OR MTCY130.18.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Sqares R., Sulston J.E.,
RA   Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: STRONG, TO M.LEPRAE B1549_C2_208.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z73902; CAA98097.1; -.
DR   MEROPS; T04.001; -.
DR   TUBERCULIST; RV1333; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM     25     45       POTENTIAL.
FT   TRANSMEM     68     88       POTENTIAL.
FT   TRANSMEM    104    124       POTENTIAL.
FT   TRANSMEM    133    153       POTENTIAL.
FT   TRANSMEM    161    181       POTENTIAL.
FT   TRANSMEM    224    244       POTENTIAL.
FT   TRANSMEM    276    296       POTENTIAL.
FT   TRANSMEM    302    322       POTENTIAL.
SQ   SEQUENCE   344 AA;  33952 MW;  296A055FECF8E650 CRC64;

  Query Match             13.0%;  Score 60;  DB 1;  Length 344;
  Best Local Similarity   32.5%;  Pred. No. 5.1;
  Matches   27;  Conservative    7;  Mismatches   17;  Indels   32;  Gaps    6;

Qy      23 GCIDARTG----RETAAL---NSVRLLQSLL-------------VLR-----RRPCSRD 56
           | :| | |    |||  |   |||| : :||             |:|     ||  : |
Db      43 GAVDCRGGAPGTRETDLLDPANSVRFVDALLLAGGSAYGLAAADGVMRWLEEHRRGVAMD 102

Qy      57 GSGLP-TPGAFAFHTEFIHVPVG 78
              :|  |||  |      :|||
Db     103 SGVVPIVPGAVIF-----DLPVG 120


RESULT  10
DAT_HAEIN
ID   DAT_HAEIN      STANDARD;      PRT;   454 AA.
AC   P44951;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   DIAMINOBUTYRATE--PYRUVATE AMINOTRANSFERASE (EC 2.6.1.46) (L-
DE   DIAMINOBUTYRIC ACID TRANSAMINASE) (DIAMINOBUTYRATE TRANSAMINASE)
DE   (DABA AMINOTRANSFERASE) (DABA-AT).
GN   DAT OR HI0949.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
```

```
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- CATALYTIC ACTIVITY: L-2,4-DIAMINOBUTANOATE + 2-OXOGLUTARATE = L-
CC       ASPARTATE 4-SEMIALDEHYDE + L-GLUTAMATE.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC       AMINOTRANSFERASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32776; AAC22610.1; -.
DR   HSSP; P16932; 1DGE.
DR   TIGR; HI0949; -.
DR   INTERPRO; IPR000954; -.
DR   PFAM; PF00202; aminotran_3; 1.
DR   PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
KW   Transferase; Aminotransferase; Pyridoxal phosphate.
FT   BINDING     287    287       PYRIDOXAL PHOSPHATE (POTENTIAL).
SQ   SEQUENCE   454 AA;  49368 MW;  B4B26F3AB7C55063 CRC64;


  Query Match             13.0%;  Score 60;  DB 1;  Length 454;
  Best Local Similarity   30.6%;  Pred. No. 6.8;
  Matches   26;  Conservative    8;  Mismatches   21;  Indels   30;  Gaps    5;

Qy      10 YPQKLAFAECLCRGC----------IDARTGRETAAL--NSVRLLQSLLVLRRRPCSRD- 56
           ||:|| ||    :||          :|   |  | ||  |   |:|::         :|
Db      31 YPRKLPFAYAKAQGCWVTDVEGNEYLDFLAGAGTLALGHNHPILMQAI---------KDV 81

Qy      57 -GSGLP-------TPGAFAFHTEFI 73
            ||||        ||   ||  | :
Db      82 LDSGLPLHTLDLTTPLKDAFSELL 106


RESULT  11
CLAT_PIG
ID   CLAT_PIG       STANDARD;      PRT;   640 AA.
AC   P13222;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CHOLINE O-ACETYLTRANSFERASE (EC 2.3.1.6) (CHOACTASE) (CHOLINE
DE   ACETYLASE).
GN   CHAT.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENTRAL SPINAL CORD;
RX   MEDLINE=88097472; PubMed=3480542;
RA   Berrard S., Brice A., Lottspeich F., Braun A., Barde Y.-A., Mallet J.;
RT   "cDNA cloning and complete sequence of porcine choline
RT   acetyltransferase: in vitro translation of the corresponding RNA
RT   yields an active protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:9280-9284(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENTRAL SPINAL CORD;
RX   MEDLINE=89229974; PubMed=2713713;
RA   Berrard S., Brice A., Mallet J.;
RT   "Molecular genetic approach to the study of mammalian choline
RT   acetyltransferase.";
RL   Brain Res. Bull. 22:147-153(1989).
RN   [3]
RP   SEQUENCE OF 1-11.
RC   TISSUE=BRAIN;
RX   MEDLINE=87085562; PubMed=3794697;
RA   Braun A., Barde Y.-A., Lottspeich F., Mewes H.-W., Thoenen H.;
RT   "N-terminal sequence of pig brain choline acetyltransferase purified
RT   by a rapid procedure.";
RL   J. Neurochem. 48:16-21(1987).
CC   -!- FUNCTION: CATALYZES THE BIOSYNTHESIS OF THE NEUROTRANSMITTER
CC       ACETYLCHOLINE.
CC   -!- CATALYTIC ACTIVITY: ACETYL-COA + CHOLINE = COA + O-ACETYLCHOLINE.
CC   -!- SIMILARITY: BELONGS TO THE CARNITINE/CHOLINE ACETYLTRANSFERASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03021; AAA31000.1; -.
DR   EMBL; M27736; AAA31015.1; -.
DR   PIR; A28047; A28047.
DR   PIR; A39961; A39961.
DR   INTERPRO; IPR000542; -.
DR   PFAM; PF00755; Carn_acyltransf; 1.
DR   PROSITE; PS00439; ACYLTRANSF_C_1; 1.
DR   PROSITE; PS00440; ACYLTRANSF_C_2; 1.
KW   Transferase; Acyltransferase; Neurotransmitter biosynthesis.
FT   INIT_MET      0      0
FT   ACT_SITE    333    333       POTENTIAL.
SQ   SEQUENCE   640 AA;  71599 MW;  5ECC27BE8B7CC317 CRC64;


  Query Match             12.6%;  Score 58.5;  DB 1;  Length 640;
  Best Local Similarity   32.5%;  Pred. No. 15;
  Matches   13;  Conservative    9;  Mismatches   13;  Indels    5;  Gaps    1;

Qy      19 CLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGS 58
           |:|  |:||  | | :  |     ::| :|    ||::|:
Db     276 CICLVCLDAPGGMELSDTN-----RALQLLHGGGCSKNGA 310


RESULT  12
CLPX_HELPJ
ID   CLPX_HELPJ     STANDARD;      PRT;   452 AA.
AC   Q9ZJL8;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPX.
GN   CLPX OR HP1374.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- FUNCTION: ATP-DEPENDENT SPECIFICITY COMPONENT OF THE CLP PROTEASE.
CC       IT DIRECTS THE PROTEASE TO SPECIFIC SUBSTRATES. CAN PERFORM
CC       CHAPERONE FUNCTIONS IN THE ABSENCE OF CLPP (BY SIMILARITY).
CC   -!- SUBUNIT: HETERODIMER OF CLPP AND CLPX (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE001551; AAD06862.1; -.
DR   INTERPRO; IPR001939; -.
DR   PFAM; PF00004; AAA; 1.
KW   Chaperone; ATP-binding.
FT   ZN_FING      7     39       C4-TYPE.
FT   NP_BIND    155    162       ATP (POTENTIAL).
SQ   SEQUENCE   452 AA;  51166 MW;  3C1D4A719296474B CRC64;


  Query Match             12.4%;  Score 57.5;  DB 1;  Length 452;
  Best Local Similarity   29.3%;  Pred. No. 14;
  Matches   22;  Conservative    8;  Mismatches   30;  Indels   15;  Gaps    3;

Qy       7 EDRYPQK--LAFAE------CLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGS 58
           | | |:|  : ||       |:|  |||  '|      :    : ||| |:|    |  |
Db      14 ESRDPKKRRIIFASNLNKDVCVCEYCIDVMHGELHKYDSKYDRMDSLLALKRDRLRRMES 73

Qy      59 GLPTPGAFAFHTEFI 73
                  |:  ||:
Db      74 S-------AYEEEFL 81


RESULT  13
YJ85_MYCTU
ID   YJ85_MYCTU     STANDARD;      PRT;   303 AA.
AC   Q10872;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL TRANSCRIPTIONAL REGULATOR RV1985C.
GN   RV1985C OR MTCY39.34.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Sqares R., Sulston J.E.,
RA   Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z74025; CAA98410.1; -.
DR   TUBERCULIST; RV1985C; -.
DR   INTERPRO; IPR000847; -.
DR   PFAM; PF00126; HTH_1; 1.
DR   PRINTS; PR00039; HTHLYSR.
DR   PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW   Hypothetical protein; Transcription regulation; DNA-binding.
FT   DNA_BIND     23     42       H-T-H MOTIF (POTENTIAL).
SQ   SEQUENCE   303 AA;  32836 MW;  344BDB1056207144 CRC64;


  Query Match             12.3%;  Score 57;  DB 1;  Length 303;
  Best Local Similarity   28.8%;  Pred. No. 10;
  Matches   17;  Conservative   10;  Mismatches   24;  Indels    8;  Gaps    2;

Qy      11 PQKLAFAECLCRGCIDARTGR----ETAALNSVRLLQ----SLLVLRRRPCSRDGSGLP 61
           || | |  :  |  ||   |      :|    :: |:     :||:| :||    :|:|
Db       9 PQLAALAAVVELGSFDAAAERLHVTPSAVSQRIKSLEQQVGQVLVVREKPCRATTAGIP 67


RESULT  14
AROA_AGRSP
ID   AROA_AGRSP     STANDARD;      PRT;   455 AA.
AC   Q9R4E4;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-
DE   ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE) (EPSPS).
GN   AROA.
OS   Agrobacterium sp. (strain CP4).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Agrobacterium.
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-28; 47-61 AND 321-333.
RA   Barry G.F., Kishore G.M., Padgette S.R., Stallings W.C.;
RT   "Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";
RL   Patent number US5633435, 27-MAY-1997.
RN   [2]
RP   SEQUENCE OF 1-15.
RX   MEDLINE=96182485; PubMed=8598558;
RA   Harrison L.A., Bailey M.R., Naylor M.W., Ream J.E., Hammond B.G.,
RA   Nida D.L., Burnette B.L., Nickson T.E., Mitsky T.A., Taylor M.L.,
RA   Fuchs R.L., Padgette S.R.;
RT   "The expressed protein in glyphosate-tolerant soybean, 5-
RT   enolpyruvylshikimate-3-phosphate synthase from Agrobacterium sp.
RT   strain CP4, is rapidly digested in vitro and is not toxic to acutely
RT   gavaged mice.";
RL   J. Nutr. 126:728-740(1996).
CC   -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE =
CC       ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC   -!- BIOTECHNOLOGY: INTRODUCED BY GENETIC MANIPULATION AND EXPRESSED IN
CC       GLYPHOSATE-TOLERANT SOYBEAN, CANOLA, COTTON AND MAIZE BY MONSANTO.
CC       DEVELOPED TO PROVIDE NEW WEED-CONTROL OPTIONS FOR FARMERS.
CC       EXPRESSION OF THIS PROTEIN IN PLANTS IMPARTS HIGH LEVELS OF
CC       GLYPHOSATE TOLERANCE.
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Herbicide resistance;
KW   Genetically modified food.
FT   CONFLICT      2      2       S -> L (IN REF. 2).
SQ   SEQUENCE   455 AA;  47588 MW;  236580D08D6EF422 CRC64;


  Query Match             12.2%;  Score 56.5;  DB 1;  Length 455;
  Best Local Similarity   28.7%;  Pred. No. 18;
  Matches   27;  Conservative   10;  Mismatches   32;  Indels   25;  Gaps    5;

Qy       3 VDTDEDRYPQKL----------AFAE--CLCRGCIDAR---TGRETAALNSVRLL----- 42
           |   ||| |  :           ||||   :  |  :|   :  | :|  | ::|
Db     314 VTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCD 373

Qy      43 --QSLLVLRRRPCSRDGSGLPTPGAFAFHTEFIH 74
             :: ||:| ||   || ||      |  |   |
Db     374 EGETSLVVRGRP---DGKGLGNASGAAVATHLDH 404
```

```
RESULT 15
CAR4_RHINI
ID   CAR4_RHINI     STANDARD;      PRT;   398 AA.
AC   Q03700;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   RHIZOPUSPEPSIN 4 PRECURSOR (EC 3.4.23.21) (ASPARTATE PROTEASE).
OS   Rhizopus niveus.
OC   Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC   Rhizopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAMAZAKI / IFO 4810;
RA   Horiuchi H., Nakamura H., Okazaki T., Yano K., Takagi M.;
RL   Submitted (DEC-1990) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS WITH BROAD SPECIFICITY
CC       SIMILAR TO THAT OF PEPSIN A, PREFERRING HYDROPHOBIC RESIDUES AT P1
CC       AND P1'. CLOTS MILK AND ACTIVATES TRYPSINOGEN. DOES NOT CLEAVE
CC       4-GLN-|-HIS-5, BUT DOES CLEAVE 10-HIS-|-LEU-11 AND 12-VAL-|-GLU-13
CC       IN B CHAIN OF INSULIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC       EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X56992; CAA40309.1; -.
DR   HSSP; P06026; 4APR.
DR   MEROPS; A01.012; -.
DR   INTERPRO; IPR001461; -.
DR   INTERPRO; IPR001969; -.
DR   PFAM; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Zymogen; Signal; Multigene family.
FT   SIGNAL        1      ?       POTENTIAL.
FT   PROPEP        ?     74       POTENTIAL.
FT   CHAIN        75    398       RHIZOPUSPEPSIN 4.
FT   ACT_SITE    108    108       BY SIMILARITY.
FT   ACT_SITE    291    291       BY SIMILARITY.
FT   DISULFID    121    124       BY SIMILARITY.
FT   DISULFID    325    358       BY SIMILARITY.
FT   DOMAIN       58     67       POLY-SER.
FT   DOMAIN      237    240       POLY-GLY.
SQ   SEQUENCE   398 AA;  41408 MW;  70CA63FEFB5C05D5 CRC64;


  Query Match             12.1%;  Score 56;  DB 1;  Length 398;
  Best Local Similarity   22.3%;  Pred. No. 18;
  Matches   21;  Conservative   11;  Mismatches   30;  Indels   32;  Gaps    3;

Qy       2 RVDTDEDRYPQKLAFAECLCRGC----------------IDART--------------G 30
           ::  | |     | ||  ||  |                  | ||              |
Db     102 KLKLDFDTGSSDLWFASTLCTNCGSSQTKYDPSQSSTYAKDGRTWSISYGDGSSASGILG 161

Qy      31 RETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPG 64
           ::|  |  :::   :: | :|  |   || |: |
Db     162 KDTVNLGGLKIKNQIIELAKREASSFSSG-PSDG 194


Search completed: February 16, 2001, 12:52:57
Job time: 187 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:48:14 ; Search time 23.57 Seconds
(without alignments)
250.630 Million cell updates/sec

Title: US-09-320-713-2
Perfect score: 463
Sequence: 1 ARVDTDEDRYPQKLAFAECL.....FHTEFIHVPVGCTCVLPRSV 87

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	124	26.8	151	1	B45351	immediate-early pr
2	119	25.7	147	2	JC4628	cytotoxic T-lympho
3	119	25.7	150	2	I49623	cytotoxic T-lympho
4	88	19.0	148	2	T21334	hypothetical prote
5	67	14.5	146	2	T36476	hypothetical prote
6	63.5	13.7	563	2	T04359	pectin methylester
7	63	13.6	362	2	S22395	fetuin precursor -
8	62	13.4	1737	2	T00209	MEGF8 protein - hu
9	61	13.2	2222	2	T13924	sdk protein - frui
10	60.5	13.1	306	2	S50689	hypothetical prote
11	60	13.0	227	2	I39313	zinc-finger protei
12	60	13.0	245	2	T27780	hypothetical prote
13	60	13.0	344	2	A70771	probable hydrolase
14	60	13.0	436	2	T36490	probable transmemb
15	60	13.0	454	2	C64104	gamma-aminobutyrat
16	59.5	12.9	491	2	T19635	hypothetical prote
17	59.5	12.9	581	2	T04844	probable serine/th
18	58.5	12.6	102	2	T36634	probable transcrip
19	58.5	12.6	641	2	A39961	choline O-acetyltr
20	58.5	12.6	907	2	T02417	probable finger pr
21	58	12.5	183	2	A75605	hypothetical prote
22	58	12.5	194	2	T00941	hypothetical prote
23	58	12.5	524	2	E75574	aldehyde dehydroge
24	58	12.5	4767	2	T31345	hypothetical prote
25	57.5	12.4	452	2	F71826	ATP-dependent prot
26	57.5	12.4	877	2	T43449	hypothetical prote
27	57.5	12.4	1067	2	T04274	hypothetical prote
28	57	12.3	303	2	G70756	hypothetical prote
29	56.5	12.2	430	2	T12541	hypothetical prote
30	56	12.1	204	2	T28939	hypothetical prote
31	56	12.1	1964	2	T09059	notch4 - mouse
32	55.5	12.0	298	2	B42643	tRNA isopentenyltr
33	55.5	12.0	328	2	A69855	low-affinity inorg
34	55.5	12.0	363	2	I50475	dopamine D1 recept
35	55.5	12.0	375	2	JC7287	G-protein coupled
36	55.5	12.0	445	2	T43804	gamma-aminobutyrat
37	55.5	12.0	489	2	S68280	protein disulfide-
38	55.5	12.0	547	2	S19607	alkaline phosphata
39	55.5	12.0	554	2	T33894	hypothetical prote
40	55.5	12.0	597	2	T21883	hypothetical prote
41	55	11.9	297	2	JL0032	hypothetical 31.7K
42	55	11.9	350	2	T34829	transcription regu
43	55	11.9	456	2	S76369	hypothetical prote
44	55	11.9	527	1	SAHUP	saposin precursor
45	55	11.9	634	2	A83386	hypothetical prote

ALIGNMENTS

RESULT 1
B45351
immediate-early protein 2 - saimiriine herpesvirus 1 (strain 11)
N;Alternate names: hypothetical protein ORF13
C;Species: saimiriine herpesvirus 1
A;Note: host Saimiri sciureus (common squirrel monkey)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C;Accession: B45351; D36807
R;Nicholas, J.; Smith, E.P.; Coles, L.; Honess, R.
Virology 179, 189-200, 1990
A;Title: Gene expression in cells infected with gammaherpesvirus saimiri: properties
A;Reference number: A45351; MUID:91021021
A;Accession: B45351
A;Molecule type: mRNA
A;Residues: 1-151 <NIC>
A;Cross-references: GB:M60286; NID:g331040; PIDN:AAA46156.1; PID:g331042
R;Albrecht, J.
submitted to the EMBL Data Library, January 1992
A;Description: Primary structure of the herpesvirus saimiri genome.
A;Reference number: A36806
A;Accession: D36807
A;Molecule type: DNA
A;Residues: 1-151 <ALB>
A;Cross-references: GB:X64346; NID:g60320; PIDN:CAA45636.1; PID:g60334
R;Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.;
J. Virol. 66, 5047-5058, 1992
A;Title: Primary structure of the herpesvirus saimiri genome.
A;Reference number: A37309; MUID:92333688
A;Contents: annotation; protein-coding frames
A;Note: neither protein nor nucleotide sequence is given
C;Genetics:
A;Gene: 13
C;Superfamily: saimiri herpesvirus immediate-early protein 2
C;Keywords: early protein

Query Match 26.8%; Score 124; DB 1; Length 151;
Best Local Similarity 37.0%; Pred. No. 1.6e-07;
Matches 30; Conservative 13; Mismatches 24; Indels 14; Gaps 3;

```
Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:||||  :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||| |
Db     128 ---SFRLEKMLVTVGCTCVTP 145
```

RESULT 2
JC4628
cytotoxic T-lymphocyte-associated antigen 8 precursor - mouse

```
N;Alternate names: CTLA8 protein
C;Species: Mus musculus (house mouse)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 26-Aug-1999
C;Accession: JC4628
R;Yao, Z.; Timour, M.; Painter, S.; Fanslow, W.; Spriggs, M.
Gene 168, 223-225, 1996
A;Title: Complete nucleotide sequence of the mouse CTLA8 gene.
A;Reference number: JC4628; MUID:96194901
A;Accession: JC4628
A;Molecule type: DNA
A;Residues: 1-147 <YAO>
A;Cross-references: GB:U35108; NID:g1244499; PIDN:AAA93253.1; PID:g1244500
C;Genetics:
A;Gene: ctla8
A;Introns: 69/2
C;Superfamily: saimiri herpesvirus immediate-early protein 2
C;Keywords: cytokine; glycoprotein; lymphocyte
F;1-14/Domain: signal sequence #status predicted <SIG>
F;15-147/Product: cytotoxic T-lymphocyte-associated antigen 8 #status predicted <MAT>
F;60/Binding site: carbohydrate (Asn) (covalent) #status predicted


  Query Match             25.7%;  Score 119;  DB 2;  Length 147;
  Best Local Similarity   39.0%;  Pred. No. 6.1e-07;
  Matches   30;  Conservative   11;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGA 65
           | ||||  :  |:|  : |::|  |:    :||| : | :|||:| | |
Db      73 DPDRYPSVIWEAQCRHQRCVNAE-GKLDHHMNSVLIQQEILVLKREPES---------CP 122

Qy      66 FAFHTEFIHVPVGCTCV 82
           | |  | : | ||||||
Db     123 FTFRVEKMLVGVGCTCV 139


RESULT   3
I49623
cytotoxic T-lymphocyte-associated antigen 8 precursor - mouse
N;Alternate names: immediate-early protein 2 (ORF13) homolog
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 26-Aug-1999
C;Accession: I49623
R;Rouvier, E.; Luciani, M.
J. Immunol. 150, 5445-5456, 1993
A;Title: CTLA-8, cloned from an activated T cell, bearing AU-rich messenger RNA instabil
A;Reference number: I49623; MUID:93294300
A;Accession: I49623
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-150 <RES>
A;Cross-references: GB:L13839; NID:g293329; PIDN:AAA37490.1; PID:g293330
C;Genetics:
A;Gene: Ctla-8
C;Superfamily: saimiri herpesvirus immediate-early protein 2


  Query Match             25.7%;  Score 119;  DB 2;  Length 150;
  Best Local Similarity   38.0%;  Pred. No. 6.3e-07;
  Matches   30;  Conservative   11;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRP--CSRDGSGLPTP 63
           | ||||  :  |:|  : |::|  |:    :||| : | :|||:| |  |
Db      76 DPDRYPSVIWEAQCRHQRCVNAE-GKLDHHMNSVLIQQEILVLKREPEKC---------- 124

Qy      64 GAFAFHTEFIHVPVGCTCV 82
             | |  | : | ||||||
Db     125 -PFTFRVEKMLVGVGCTCV 142


RESULT   4
T21334
hypothetical protein F25D1.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21334
R;Kelly, P.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19408
A;Accession: T21334
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-148 <WIL>
A;Cross-references: EMBL:Z73973; PIDN:CAA98268.1; GSPDB:GN00023; CESP:F25D1.3
A;Experimental source: clone F25D1
C;Genetics:
A;Gene: CESP:F25D1.3
A;Map position: 5
A;Introns: 60/2


  Query Match             19.0%;  Score 88;  DB 2;  Length 148;
  Best Local Similarity   32.6%;  Pred. No. 0.0031;
  Matches   28;  Conservative   15;  Mismatches   31;  Indels   12;  Gaps    4;

Qy       1 ARVDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGL 60
             :||:  | | |: :| : |||      |  | :     : ::: : :|||  | |  :||
Db      59 SRVNYQESREPKLIAESVCLC------RKSRGSTGAFCMPIVRKVPILRRVSCDR-STGL 111

Qy      61 PTPGAFAFHTEFIHVPVGCTCVLPRS 86
                :  || |  |||  ||||:
Db     112 ---WNYVRSTELI--TVGCHSVLPRT 132


RESULT   5
T36476
hypothetical protein SCF85.16 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T36476
R;Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21607
A;Accession: T36476
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-146 <SEE>
A;Cross-references: EMBL:AL110470; PIDN:CAB54173.1; GSPDB:GN00070; SCOEDB:SCF85.16
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCF85.16


  Query Match             14.5%;  Score 67;  DB 2;  Length 146;
  Best Local Similarity   43.1%;  Pred. No. 0.97;
  Matches   22;  Conservative    2;  Mismatches   15;  Indels   12;  Gaps    3;

Qy      25 IDARTGRETAALNSVRLLQSLLVLRRRPCSRDGS-GLP---TPGAFAFHTE 71
           : ||| :   ||  ||         ||| ||||| | |    || |   ||
Db     102 LQARTPQTAQALVGVR--------ARRPLSRDGSAGAPRGRRPGLFLTRTE 144


RESULT   6
T04359
pectin methylesterase-like protein - maize
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C;Accession: T04359
R;Wakeley, P.R.; Rogers, H.J.; Rozycka, M.; Greenland, A.J.; Hussey, P.J.
Plant Mol. Biol. 37, 187-192, 1998
A;Title: A maize pectin methylesterase-like gene, ZmC5, specifically expressed in pol
A;Reference number: Z15309; MUID:98281586
A;Accession: T04359
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-563 <WAK>
A;Cross-references: EMBL:Y13285; NID:g2239261; PIDN:CAA73733.1; PID:g2239262
```

```
A;Experimental source: strain A188; pollen
C;Genetics:
A;Gene: C5
C;Superfamily: pectinesterase


  Query Match             13.7%;  Score 63.5;  DB 2;  Length 563;
  Best Local Similarity   27.3%;  Pred. No. 9.9;
  Matches   18;  Conservative   10;  Mismatches   29;  Indels    9;  Gaps    2;

Qy       2 RVDTDEDR---YPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGS 58
           | |  :|     |::  |  |:  | ||   |      ||  : |: |:: |||     :
Db     373 RFDAFQDTLYVQPRRQFFRNCVVSGTIDFIFG------NSAAVFQNCLIITRRPMDNQQN 426

Qy      59 GLPTPG 64
             :  |
Db     427 SVTAHG 432


RESULT   7
S22395
fetuin precursor - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jul-2000
C;Accession: S22395; S22142
R;Brown, W.M.; Dziegielewska, K.M.; Saunders, N.R.; Christie, D.L.; Nawratil, P.; Muelle
Eur. J. Biochem. 205, 321-331, 1992
A;Title: The nucleotide and deduced amino acid structures of sheep and pig fetuin. Comm
A;Reference number: S22394; MUID:92209519
A;Accession: S22395
A;Molecule type: mRNA
A;Residues: 1-362 <BRW>
A;Cross-references: EMBL:X56021; NID:g2104; PIDN:CAA39498.1; PID:g3980229
C;Superfamily: alpha-2-HS-glycoprotein; cystatin homology
C;Keywords: calcium binding; EF hand; glycoprotein
F;1-15/Domain: signal sequence (fragment) #status predicted <SIG>
F;16-362/Product: fetuin #status predicted <MAT>
F;20-134/Domain: cystatin homology <CY1>
F;143-249/Domain: cystatin homology <CY2>
F;96,153,173/Binding site: carbohydrate (Asn) (covalent) #status predicted


  Query Match             13.6%;  Score 63;  DB 2;  Length 362;
  Best Local Similarity   25.4%;  Pred. No. 7.3;
  Matches   17;  Conservative    8;  Mismatches   42;  Indels    0;  Gaps    0;

Qy      21 CRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGAFAFHTEFIHVPVGCT 80
           |:| : |:   |  |:        :||: :|   |    |  | |  :   ||
Db     227 CKGTVTAKVNEEDVAVTCTVFQTQPVVLQPQPAGADAGATPVVDAAATASPLADVPAASL 286

Qy      81 CVLPRSV 87
             | |  |
Db     287 VVGPMVV 293


RESULT   8
T00209
MEGF8 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00209
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A;Reference number: Z14126; MUID:98360089
A;Accession: T00209
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1737 <NAK>
A;Cross-references: EMBL:AB011541; NID:g3449307; PIDN:BAA32469.1; PID:g3449308
A;Experimental source: brain; clone HG1392
C;Genetics:
A;Gene: MEGF8
A;Map position: 19q12


  Query Match             13.4%;  Score 62;  DB 2;  Length 1737;
  Best Local Similarity   27.8%;  Pred. No. 47;
  Matches   20;  Conservative    7;  Mismatches   33;  Indels   12;  Gaps    2;

Qy      19 CLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPC------SRDGSGLPTPGAFAFHTEF 72
           |:  ||     | |: :|   :  |  | |||  |       :|| |    |  :   :
Db     994 CMAGGCGRLLRGPESCSLGCAQATQCALCLRRPHCGWCAWGGQDGGGRCMEGGLSGPRD- 1052

Qy      73 IHVPVGCTCVLP 84
                | ||  |
Db    1053 -----GLTCGRP 1059


RESULT   9
T13924
sdk protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T13924
R;Nguyen, D.N.; Liu, Y.; Litsky, M.L.; Reinke, R.
submitted to the EMBL Data Library, February 1997
A;Description: Sidekick, a member of the immunoglobulin superfamily, is required for
A;Reference number: Z17809
A;Accession: T13924
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2222 <NGU>
A;Cross-references: EMBL:U88578; NID:g4099554; PID:g4099555; PIDN:AAD09632.1
C;Genetics:
A;Gene: sdk


  Query Match             13.2%;  Score 61;  DB 2;  Length 2222;
  Best Local Similarity   34.5%;  Pred. No. 79;
  Matches   19;  Conservative    9;  Mismatches   17;  Indels   10;  Gaps    2;

Qy      27 ARTGRETAALNSVRLLQS--LLVLRRRPCSRDGSGL--------PTPGAFAFHTE 71
           :| |  |  ||||  |:|  |  | |:  ||   |:         |:| : |:|::
Db    2058 SRHGVGTGTLNSVGTLRSGTLGTLGRKSTSRPPPGVHLGKSPPRPSPASVAYHSD 2112


RESULT  10
S50689
hypothetical protein YER186c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 29-Oct-1999
C;Accession: S50689
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmids 9163 and 9132.
A;Reference number: S50428
A;Accession: S50689
A;Molecule type: DNA
A;Residues: 1-306 <DIE>
A;Cross-references: EMBL:U18922; NID:g603405; PID:g603427; GSPDB:GN00005; MIPS:YER186
C;Genetics:
A;Gene: MIPS:YER186c
A;Map position: 5R


  Query Match             13.1%;  Score 60.5;  DB 2;  Length 306;
  Best Local Similarity   31.0%;  Pred. No. 12;
  Matches   22;  Conservative    9;  Mismatches   31;  Indels    9;  Gaps    4;

Qy       6 DEDR--YPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRP---CSRDGSGL 60
           :|||   | |  |      | :  |  ::     | | : :|   |: |    : || ||
Db      22 NEDRKDVPDKFKFDTVTKTGMVKLRVFKDDLIFKSQRSI-NLFASRKHPFKSFTADGEGL 80

Qy      61 PTPGAFAFHTE 71
           |    ||| |:
```

```
Db      81 P---LFAFRTK 88


RESULT  11
I39313
zinc-finger protein (ZNFpT1) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999
C;Accession: I39313
R;Huebner, K.; Druck, T.; LaForgia, S.; Lasota, J.; Croce, C.M.; Lanfrancone, L.; Donti,
Hum. Genet. 91, 217-222, 1993
A;Title: Chromosomal localization of four human zinc finger cDNAs.
A;Reference number: I39313; MUID:93239177
A;Accession: I39313
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-227 <RES>
A;Cross-references: EMBL:X65230; NID:g505543; PIDN:CAA46337.1; PID:g505544


  Query Match             13.0%;  Score 60;  DB 2;  Length 227;
  Best Local Similarity   29.7%;  Pred. No. 10;
  Matches   19;  Conservative    6;  Mismatches   31;  Indels    8;  Gaps    2;

Qy      11 PQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGAFAFHT 70
           | |    |  |  | |:||        ||| ::||   :   | : |     | :|  |
Db      87 PMKTQTVEKPC-NCQDSRTA-------SVRYVKSLSSKKSYECQKCGKAFICPSSFRGHV 138

Qy      71 EFIH 74
             |
Db     139 NSHH 142


RESULT  12
T27780
hypothetical protein ZK218.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 26-May-2000
C;Accession: T27780
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z20418
A;Accession: T27780
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-245 <WIL>
A;Cross-references: EMBL:Z82085; PIDN:CAB04984.1; GSPDB:GN00023; CESP:ZK218.1
A;Experimental source: clone ZK218
C;Genetics:
A;Gene: CESP:ZK218.1
A;Map position: 5
A;Introns: 35/3; 124/1
C;Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3


  Query Match             13.0%;  Score 60;  DB 2;  Length 245;
  Best Local Similarity   29.7%;  Pred. No. 11;
  Matches   19;  Conservative    8;  Mismatches   21;  Indels   16;  Gaps    3;

Qy      19 CLCRGCIDARTG--RETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGAFAFHTEFIHVP 76
           |   ||:||  |   : :  |:| :   : :  :| |||  |  | ||
Db     157 CNAGGCVDAIMGCANDISICNNVGMQDFVNLNCQRTCSRCSSA-PVPG------------ 203

Qy      77 VGCT 80
             |||
Db     204 -GCT 206


RESULT  13
A70771
probable hydrolase - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: A70771
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A;Reference number: A70500; MUID:98295987
A;Accession: A70771
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-344 <COL>
A;Cross-references: GB:Z73902; GB:AL123456; NID:g3261576; PIDN:CAA98097.1; PID:g4502
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv1333


  Query Match             13.0%;  Score 60;  DB 2;  Length 344;
  Best Local Similarity   32.5%;  Pred. No. 16;
  Matches   27;  Conservative    7;  Mismatches   17;  Indels   32;  Gaps    6;

Qy      23 GCIDARTG----RETAAL---NSVRLLQSLL--------------VLR-----RRPCSRD 56
           | :| | |    |||  |   |||| : :||              |:|     ||  : |
Db      43 GAVDCRGGAPGTRETDLLDPANSVRFVDALLLAGGSAYGLAAADGVMRWLEEHRRGVAMD 102

Qy      57 GSGLP-TPGAFAFHTEFIHVPVG 78
             :|  |||  |      :|||
Db     103 SGVVPIVPGAVIF-----DLPVG 120


RESULT  14
T36490
probable transmembrane transport protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: T36490
R;Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21608
A;Accession: T36490
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-436 <SAU>
A;Cross-references: EMBL:AL096822; PIDN:CAB46932.1; GSPDB:GN00070; SCOEDB:SCGD3.10c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCGD3.10c
C;Superfamily: bicyclomycin resistance protein


  Query Match             13.0%;  Score 60;  DB 2;  Length 436;
  Best Local Similarity   36.2%;  Pred. No. 20;
  Matches   21;  Conservative    7;  Mismatches   20;  Indels   10;  Gaps    2;

Qy      25 IDARTGRETAALNSVRLLQSL-----LVLRRRPCSRDGSGLPTPGAFAFHTEFIHVPV 77
           :||     | | :|| :||:|     | ||:|   |:   |      |   | :||||
Db     384 LDATGDDYTVAFSSVFVLQALGVTQILRLRKRAALRERERL-----VASRVETVHVPV 436


RESULT  15
C64104
gamma-aminobutyrate aminotransferase (EC 2.6.1.-) - Haemophilus influenzae (strain Rd
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
C;Accession: C64104; JC5879
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630
```

A;Accession: C64104
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-454 <TIGR>
A;Cross-references: GB:U32776; GB:L42023; NID:g1573969; PIDN:AAC22610.1; PID:g1573974; T
R;Ikai, H.; Yamamoto, S.
Biol. Pharm. Bull. 21, 170-173, 1998
A;Title: Two genes involved in the 1,3-diaminopropane production pathway in Haemophilus
A;Reference number: JC5879; MUID:98173550
A;Accession: JC5879
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-454 <IKA>
A;Cross-references: GB:U32776; NID:g1573969; PIDN:AAC22610.1; PID:g1573974
C;Comment: This enzyme is involved in the pathway of 1,3-diaminopropane synthesis.
C;Genetics:
A;Note: neither the complete nucleic acid sequence nor the complete translation are show
C;Superfamily: beta-alanine--pyruvate transaminase
C;Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F;287/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 13.0%; Score 60; DB 2; Length 454;
Best Local Similarity 30.6%; Pred. No. 21;
Matches 26; Conservative 8; Mismatches 21; Indels 30; Gaps 5;

```
Qy      10 YPQKLAFAECLCRGC----------IDARTGRETAAL--NSVRLLQSLLVLRRRPCSRD- 56
           ||:|| ||    :||          :|   |  | ||  |   |:|::         :|
Db      31 YPRKLPFAYAKAQGCWVTDVEGNEYLDFLAGAGTLALGHNHPILMQAI---------KDV 81

Qy      57 -GSGLP-------TPGAFAFHTEFI 73
            ||||       ||   ||  | :
Db      82 LDSGLPLHTLDLTTPLKDAFSEELL 106
```

Search completed: February 16, 2001, 12:50:39
Job time: 145 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:48:00 ; Search time 18.64 Seconds
(without alignments)
83.812 Million cell updates/sec

Title: US-09-320-713-2
Perfect score: 463
Sequence: 1 ARVDTDEDRYPQKLAFAECL.....FHTEFIHVPVGCTCVLPRSV 87

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	136	29.4	153	1	US-08-514-014-12	Sequence 12, Appl
2	136	29.4	153	2	US-08-833-823-12	Sequence 12, Appl
3	136	29.4	163	3	US-09-034-810-2	Sequence 2, Appli
4	136	29.4	163	3	US-08-685-239-2	Sequence 2, Appli
5	124	26.8	151	2	US-08-620-694A-8	Sequence 8, Appli
6	124	26.8	151	3	US-09-034-810-6	Sequence 6, Appli
7	124	26.8	151	3	US-09-022-255-8	Sequence 8, Appli
8	124	26.8	151	3	US-09-022-696-8	Sequence 8, Appli
9	124	26.8	151	3	US-08-685-239-6	Sequence 6, Appli
10	124	26.8	151	3	US-09-022-253-8	Sequence 8, Appli
11	124	26.8	151	3	US-09-022-260-8	Sequence 8, Appli
12	119	25.7	150	3	US-09-034-810-4	Sequence 4, Appli
13	119	25.7	150	3	US-08-685-239-4	Sequence 4, Appli
14	119	25.7	158	2	US-08-620-694A-7	Sequence 7, Appli
15	119	25.7	158	3	US-09-022-255-7	Sequence 7, Appli
16	119	25.7	158	3	US-09-022-696-7	Sequence 7, Appli
17	119	25.7	158	3	US-09-022-253-7	Sequence 7, Appli
18	119	25.7	158	3	US-09-022-260-7	Sequence 7, Appli
19	63	13.6	361	2	US-08-483-926A-9	Sequence 9, Appli
20	57.5	12.4	1349	3	US-08-938-291A-6	Sequence 6, Appli
21	57	12.3	362	2	US-08-737-045-13	Sequence 13, Appl
22	56.5	12.2	455	1	US-08-476-008-3	Sequence 3, Appli
23	56.5	12.2	455	1	US-08-306-063-3	Sequence 3, Appli
24	56.5	12.2	455	1	US-08-833-485-3	Sequence 3, Appli
25	56.5	12.2	455	4	PCT-US91-06148A-3	Sequence 3, Appli
26	55	11.9	523	1	US-08-100-247-2	Sequence 2, Appli
27	55	11.9	523	1	US-08-483-146A-2	Sequence 2, Appli
28	55	11.9	523	1	US-08-232-513A-3	Sequence 3, Appli
29	55	11.9	523	1	US-08-484-594A-2	Sequence 2, Appli
30	55	11.9	2439	3	US-09-335-409-7	Sequence 7, Appli
31	54.5	11.8	550	1	US-08-484-493-2	Sequence 2, Appli
32	54.5	11.8	550	1	US-08-484-494-2	Sequence 2, Appli
33	54.5	11.8	550	2	US-08-345-212-2	Sequence 2, Appli
34	54.5	11.8	550	3	US-09-249-003-2	Sequence 2, Appli
35	53	11.4	291	3	US-09-080-120A-7	Sequence 7, Appli
36	53	11.4	291	4	PCT-US95-08925-7	Sequence 7, Appli
37	52.5	11.3	349	3	US-08-630-172-6	Sequence 6, Appli
38	52.5	11.3	563	2	US-08-714-168-1	Sequence 1, Appli
39	52	11.2	1487	2	US-08-760-489-2	Sequence 2, Appli
40	52	11.2	1487	2	US-08-760-489-4	Sequence 4, Appli
41	51	11.0	95	2	US-08-722-349-3	Sequence 3, Appli
42	51	11.0	95	3	US-09-204-328-3	Sequence 3, Appli
43	51	11.0	741	1	US-08-277-231A-4	Sequence 4, Appli
44	51	11.0	741	2	US-08-473-750-7	Sequence 7, Appli
45	51	11.0	741	2	US-08-477-326-7	Sequence 7, Appli

ALIGNMENTS

```
RESULT 1
US-08-514-014-12
; Sequence 12, Application US/08514014
; Patent No. 5707829
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    TITLE OF INVENTION:  DNA SEQUENCES AND SECRETED PROTEINS
;    TITLE OF INVENTION:  ENCODED THEREBY
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc. -- Legal Affairs
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/514,014
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI6000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  153 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-514-014-12


  Query Match             29.4%;  Score 136;  DB 1;  Length 153;
  Best Local Similarity   39.3%;  Pred. No. 3.1e-11;
  Matches   33;  Conservative   14;  Mismatches   23;  Indels   14;  Gaps    3;

Qy       3 VDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRR--PCSRDGSGL 60
           |  | :||| ::  |:|   |||:|: |:|  ::||| : |  ||:||:   ||
Db      76 VTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQGCS------ 128
```

```
Qy      61 PTPGAFAFHTEFIHVPVGCTCVLP 84
                :|  | : | ||||||  |
Db     129 -----VSFQLEKVLVTVGCTCVTP 147


RESULT   2
US-08-833-823-12
; Sequence 12, Application US/08833823
; Patent No. 5969093
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    TITLE OF INVENTION:  DNA SEQUENCES AND SECRETED PROTEINS
;    TITLE OF INVENTION:  ENCODED THEREBY
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc. -- Legal Affairs
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/833,823
;      FILING DATE:  10-APR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/514,014
;      FILING DATE:  11-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI6000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  153 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-833-823-12


  Query Match             29.4%;  Score 136;  DB 2;  Length 153;
  Best Local Similarity   39.3%;  Pred. No. 3.1e-11;
  Matches   33;  Conservative   14;  Mismatches   23;  Indels   14;  Gaps    3;

Qy       3 VDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRR--PCSRDGSGL 60
            |  | :||| ::  |:|   |||:|: |:|  ::||| : |  ||:||:   ||
Db      76 VTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQGCS------ 128

Qy      61 PTPGAFAFHTEFIHVPVGCTCVLP 84
                :|  | : | ||||||  |
Db     129 -----VSFQLEKVLVTVGCTCVTP 147


RESULT   3
US-09-034-810-2
; Sequence 2, Application US/09034810
; Patent No. 6043344
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/034,810
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/685,239
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-034-810-2


  Query Match             29.4%;  Score 136;  DB 3;  Length 163;
  Best Local Similarity   39.3%;  Pred. No. 3.3e-11;
  Matches   33;  Conservative   14;  Mismatches   23;  Indels   14;  Gaps    3;

Qy       3 VDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRR--PCSRDGSGL 60
            |  | :||| ::  |:|   |||:|: |:|  ::||| : |  ||:||:   ||
Db      86 VTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQGCS------ 138

Qy      61 PTPGAFAFHTEFIHVPVGCTCVLP 84
                :|  | : | ||||||  |
Db     139 -----VSFQLEKVLVTVGCTCVTP 157


RESULT   4
US-08-685-239-2
; Sequence 2, Application US/08685239
; Patent No. 6074849
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,239
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-685-239-2


  Query Match             29.4%;  Score 136;  DB 3;  Length 163;
  Best Local Similarity   39.3%;  Pred. No. 3.3e-11;
  Matches   33;  Conservative   14;  Mismatches   23;  Indels   14;  Gaps    3;

Qy       3 VDTDEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRR--PCSRDGSGL 60
            |  | :||| ::  |:|   |||:|: |:|  ::||| : |  ||:||:   ||
Db      86 VTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQGCS------ 138

Qy      61 PTPGAFAFHTEFIHVPVGCTCVLP 84
              :|  | : | |||||| |
Db     139 -----VSFQLEKVLVTVGCTCVTP 157


RESULT   5
US-08-620-694A-8
; Sequence 8, Application US/08620694A
; Patent No. 5869286
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 5869286el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/620,694A
;      FILING DATE:  21 MARCH 1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/538,765
;      FILING DATE:  7 AUGUST 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpesvirus Saimiri
;      STRAIN:  ORF13
US-08-620-694A-8


  Query Match             26.8%;  Score 124;  DB 2;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 1.4e-09;
  Matches   30;  Conservative   13;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
            |:|||| :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
              :|  | : | |||||| |
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT   6
US-09-034-810-6
; Sequence 6, Application US/09034810
; Patent No. 6043344
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/034,810
;      FILING DATE:
;      CLASSIFICATION:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/685,239
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-034-810-6


  Query Match             26.8%;  Score 124;  DB 3;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 1.4e-09;
  Matches   30;  Conservative   13;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:||||  :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||||
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT   7
US-09-022-255-8
; Sequence 8, Application US/09022255
; Patent No. 6072033
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6072033el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,255
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/620,694
;      FILING DATE:  21 MARCH 1996
;      APPLICATION NUMBER:  USSN 08/538,765
;      FILING DATE:  7 AUGUST 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpesvirus Saimiri
;      STRAIN:  ORF13
US-09-022-255-8


  Query Match             26.8%;  Score 124;  DB 3;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 1.4e-09;
  Matches   30;  Conservative   13;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:||||  :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||||
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT   8
US-09-022-696-8
; Sequence 8, Application US/09022696
; Patent No. 6072037
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6072037el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,696
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/620,694
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
```

```
;       STRANDEDNESS:  not relevant
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Herpesvirus Saimiri
;       STRAIN:  ORF13
US-09-022-696-8


  Query Match             26.8%;  Score 124;  DB 3;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 1.4e-09;
  Matches   30;  Conservative   13;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:||||  :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||||  |
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT   9
US-08-685-239-6
; Sequence 6, Application US/08685239
; Patent No. 6074849
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,239
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-685-239-6


  Query Match             26.8%;  Score 124;  DB 3;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 1.4e-09;
  Matches   30;  Conservative   13;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:||||  :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||||  |
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT  10
US-09-022-253-8
; Sequence 8, Application US/09022253
; Patent No. 6096305
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6096305el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,253
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/620,694
;      FILING DATE:  21-MARCH-1996
;      APPLICATION NUMBER:  USSN 08/538,765
;      FILING DATE:  7 AUGUST 1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  No. 6096305 Relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpesvirus Saimiri
;      STRAIN:  ORF13
US-09-022-253-8


  Query Match             26.8%;  Score 124;  DB 3;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 1.4e-09;
  Matches   30;  Conservative   13;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:||||  :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||||  |
Db     128 ---SFRLEKMLVTVGCTCVTP 145
```

```
Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:|||| :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||||
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT 11
US-09-022-260-8
; Sequence 8, Application US/09022260
; Patent No. 6100235
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6100235el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,260
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/620,694
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpesvirus Saimiri
;      STRAIN:  ORF13
US-09-022-260-8


  Query Match             26.8%;  Score 124;  DB 3;  Length 151;
  Best Local Similarity   37.0%;  Pred. No. 1.4e-09;
  Matches   30;  Conservative   13;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRR--RPCSRDGSGLPTP 63
           |:|||| :  |:|   ||::|  |     :||| : | :||:|:  :||
Db      77 DQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRKGHQPCPN-------- 127

Qy      64 GAFAFHTEFIHVPVGCTCVLP 84
             :|  | : | ||||||
Db     128 ---SFRLEKMLVTVGCTCVTP 145


RESULT 12
US-09-034-810-4
; Sequence 4, Application US/09034810
; Patent No. 6043344
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/034,810
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/685,239
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-034-810-4


  Query Match             25.7%;  Score 119;  DB 3;  Length 150;
  Best Local Similarity   38.0%;  Pred. No. 6.7e-09;
  Matches   30;  Conservative   11;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRP--CSRDGSGLPTP 63
           | ||||  :  |:|  : |::|  |:    :||| : | :|||:| |  |
Db      76 DPDRYPSVIWEAQCRHQRCVNAE-GKLDHHMNSVLIQQEILVLKREPEKC---------- 124

Qy      64 GAFAFHTEFIHVPVGCTCV 82
             | |  | : | ||||||
Db     125 -PFTFRVEKMLVGVGCTCV 142


RESULT 13
US-08-685-239-4
; Sequence 4, Application US/08685239
```

```
; Patent No. 6074849
; GENERAL INFORMATION:
;   APPLICANT: Jacobs, Kenneth
;   APPLICANT: Kelleher, Kerry
;   APPLICANT:' Carlin, McKeough
;   APPLICANT: Goldman, Samuel
;   APPLICANT: Pittman, Debra
;   APPLICANT: Mi, Sha
;   APPLICANT: Neben, Steven
;   APPLICANT: Giannotti, JoAnn
;   APPLICANT: Golden'Fleet, Margaret
;   TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins
;   NUMBER OF SEQUENCES: 9
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Genetics Institute, Inc.
;     STREET: 87 CambridgePark Drive
;     CITY: Cambridge
;     STATE: Massachusetts
;     COUNTRY: USA
;     ZIP: 02140
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/685,239
;     FILING DATE:
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Brown, Scott A.
;     REGISTRATION NUMBER: 32,724
;     REFERENCE/DOCKET NUMBER: GI5262
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617) 498-8224
;     TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 150 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-685-239-4


  Query Match             25.7%;  Score 119;  DB 3;  Length 150;
  Best Local Similarity   38.0%;  Pred. No. 6.7e-09;
  Matches   30;  Conservative   11;  Mismatches   24;  Indels   14;  Gaps    3;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRP--CSRDGSGLPTP 63
           | ||||  :  |:|  : |::|  |:    :||| : | :|||:| |  |
Db      76 DPDRYPSVIWEAQCRHQRCVNAE-GKLDHHMNSVLIQQEILVLKREPEKC---------- 124

Qy      64 GAFAFHTEFIHVPVGCTCV 82
             | |  | : | ||||||
Db     125 -PFTFRVEKMLVGVGCTCV 142


RESULT 14
US-08-620-694A-7
; Sequence 7, Application US/08620694A
; Patent No. 5869286
; GENERAL INFORMATION:
;   APPLICANT: Yao, Zhengbin
;   APPLICANT: Spriggs, Melanie
;   APPLICANT: Fanslow, William
;   TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Immunex Corporation
;     STREET: 51 University Street
;     CITY: Seattle
;     STATE: WA
;     COUNTRY: USA
;     ZIP: 98101
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: Apple Power Macintosh
;     OPERATING SYSTEM: Apple Operating System 7.5.5
;     SOFTWARE: Microsoft Word for Apple, Version 6.0.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/620,694A
;     FILING DATE: 21 MARCH 1996
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: USSN 08/538,765
;     FILING DATE: 7 AUGUST 1995
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: USSN 08/410,535
;     FILING DATE: 23 MARCH 1995
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Perkins, Patricia Anne
;     REGISTRATION NUMBER: 34,695
;     REFERENCE/DOCKET NUMBER: 2617-B
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (206)587-0430
;     TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 158 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-620-694A-7


  Query Match             25.7%;  Score 119;  DB 2;  Length 158;
  Best Local Similarity   39.0%;  Pred. No. 7.2e-09;
  Matches   30;  Conservative   11;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGA 65
           | ||||  :  |:|  : |::|  |:    :||| : | :|||:| | |
Db      84 DPDRYPSVIWEAQCRHQRCVNAE-GKLDHHMNSVLIQQEILVLKREPES---------CP 133

Qy      66 FAFHTEFIHVPVGCTCV 82
           | |  | : | ||||||
Db     134 FTFRVEKMLVGVGCTCV 150


RESULT 15
US-09-022-255-7
; Sequence 7, Application US/09022255
; Patent No. 6072033
; GENERAL INFORMATION:
;   APPLICANT: Yao, Zhengbin
;   APPLICANT: Spriggs, Melanie
;   APPLICANT: Fanslow, William
;   TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Immunex Corporation
;     STREET: 51 University Street
;     CITY: Seattle
;     STATE: WA
;     COUNTRY: USA
;     ZIP: 98101
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: Apple Power Macintosh
;     OPERATING SYSTEM: Apple Operating System 7.5.5
;     SOFTWARE: Microsoft Word for Apple, Version 6.0.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/022,255
;     FILING DATE:
```

```
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  USSN 08/620,694
;       FILING DATE:  21 MARCH 1996
;       APPLICATION NUMBER:  USSN 08/538,765
;       FILING DATE:  7 AUGUST 1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  USSN 08/410,535
;       FILING DATE:  23 MARCH 1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Perkins, Patricia Anne
;       REGISTRATION NUMBER:  34,695
;       REFERENCE/DOCKET NUMBER:  2617-B
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206)587-0430
;       TELEFAX:  (206)
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;      LENGTH: 158 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-022-255-7


  Query Match             25.7%;  Score 119;  DB 3;  Length 158;
  Best Local Similarity   39.0%;  Pred. No. 7.2e-09;
  Matches   30;  Conservative   11;  Mismatches   26;  Indels   10;  Gaps    2;

Qy       6 DEDRYPQKLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGA 65
           | ||||  :  |:|  : |::|  |:    :||| : | :|||:| | |
Db      84 DPDRYPSVIWEAQCRHQRCVNAE-GKLDHHMNSVLIQQEILVLKREPES---------CP 133

Qy      66 FAFHTEFIHVPVGCTCV 82
           | |  | : | ||||||
Db     134 FTFRVEKMLVGVGCTCV 150


Search completed: February 16, 2001, 12:50:10
Job time: 130 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 20:41:30 ; Search time 2862.16 Seconds
(without alignments)
2936.015 Million cell updates/sec

Title: US-09-320-713-3
Perfect score: 1642
Sequence: 1 ggaattcggcacgagctcgt.....tgaaaaaaaaaaaaaaaaaa 1642

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_ph:*
6: gb_pl1:*
7: gb_pl2:*
8: gb_pr1:*
9: gb_pr2:*
10: gb_pr3:*
11: em_fun:*
12: em_hum1:*
13: em_hum2:*
14: em_in:*
15: em_om:*
16: em_or:*
17: em_ov:*
18: em_pat:*
19: em_ph:*
20: em_pl:*
21: em_ro:*
22: em_sts:*
23: em_sy:*
24: em_un:*
25: em_vi:*
26: gb_htg1:*
27: gb_htg2:*
28: gb_in1:*
29: gb_in2:*
30: em_ba1:*
31: em_ba2:*
32: em_hum3:*
33: em_hum4:*
34: gb_pr4:*
35: gb_htg3:*
36: gb_htg4:*
37: gb_htg5:*
38: gb_htg6:*
39: gb_htg7:*
40: em_htg1:*
41: em_htg2:*
42: em_htg3:*
43: em_hum5:*
44: gb_pl3:*
45: gb_pr5:*
46: gb_htg8:*
47: gb_htg9:*
48: gb_htg10:*
49: gb_htg11:*
50: gb_htg12:*
51: gb_htg13:*
52: gb_htg14:*
53: gb_in3:*
54: gb_htg15:*
55: gb_htg16:*
56: gb_htg17:*
57: em_htg4:*
58: em_htg5:*
59: em_htg6:*
60: em_htg7:*
61: em_hum6:*
62: gb_htg18:*
63: gb_htg19:*
64: gb_ba3:*
65: em_htg8:*
66: em_htg9:*
67: em_htg10:*
68: gb_pr6:*
69: gb_pr7:*
70: gb_htg20:*
71: gb_htg21:*
72: gb_htg22:*
73: gb_htg23:*
74: gb_ro:*
75: gb_sts1:*
76: gb_sts2:*
77: gb_sy:*
78: gb_un:*
79: gb_vi1:*
80: gb_vi2:*
81: gb_pat1:*
82: gb_pat2:*
83: em_htg0:*
84: gb_htg24:*
85: gb_pr8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	1472.2	89.7	140800	38	AC018392	AC018392 Homo sapi
c	2	1455.2	88.6	142376	70	AL161772	AL161772 Homo sapi
c	3	624.4	38.0	629	85	HUMNOTIB	L23207 Human chrom
c	4	294.4	17.9	296	85	HUMCH13C	L30108 Homo sapien
c	5	252.8	15.4	323	85	HUMNOTIA	L23206 Human chrom
c	6	245.2	14.9	250	75	G31083	G31083 human STS S
c	7	77	4.7	162005	71	AL356002	AL356002 Homo sapi
	8	73.2	4.5	687	34	AF152098	AF152098 Homo sapi
	9	73.2	4.5	688	34	AF184969	AF184969 Homo sapi
	10	73.2	4.5	711	34	AF212311	AF212311 Homo sapi
c	11	72.8	4.4	42655	64	SC7H2	AL109732 Streptomy
	12	65	4.0	1429	7	D76415	D76415 Rice seedli
	13	65	4.0	2588	6	AF099203	AF099203 Oryza sat
	14	63.6	3.9	3578	1	AF005277	AF005277 Cellulomo
	15	63.6	3.9	180630	49	AC025783	AC025783 Oryza sat
	16	63.6	3.9	224517	56	AC074335	AC074335 Mus muscu
c	17	63.4	3.9	332976	63	AC084055	AC084055 Mus muscu
c	18	62.6	3.8	144500	35	AC010231	AC010231 Homo sapi
	19	61.4	3.7	4615	64	STMWHIB12X	L22864 Streptomyce
	20	61.4	3.7	38390	64	SC2H12	AL359215 Streptomy
c	21	61	3.7	300695	62	AC079431	AC079431 Mus muscu

	22	60.8	3.7	143209	7	AP001550	AP001550 Oryza sat
c	23	60.6	3.7	8058	64	SCE56	AL138852 Streptomy
	24	60.2	3.7	33810	64	SCE6	AL353832 Streptomy
	25	60.2	3.7	77457	1	AF210249	AF210249 Streptomy
	26	60	3.7	2544	6	AB004819	AB004819 Oryza sat
c	27	59.6	3.6	2381	34	AK025816	AK025816 Homo sapi
c	28	59.6	3.6	2953	34	AK025470	AK025470 Homo sapi
	29	59.4	3.6	692	74	AF184970	AF184970 Mus muscu
c	30	59.4	3.6	14634	64	SC8B7	AL031225 Streptomy
c	31	58.8	3.6	78220	47	AC023212	AC023212 Homo sapi
c	32	58.6	3.6	30991	64	SCC53	AL357591 Streptomy
	33	58.4	3.6	150074	37	AC016007	AC016007 Homo sapi
c	34	58.2	3.5	38426	64	SCF55	AL132991 Streptomy
c	35	58.2	3.5	107109	69	HSDJ781B1	AL118522 Human DNA
c	36	57.8	3.5	165943	47	AC022205	AC022205 Homo sapi
c	37	57.4	3.5	190014	7	AP000836	AP000836 Oryza sat
c	38	57	3.5	200580	63	AC084064	AC084064 Homo sapi
	39	56.8	3.5	1047	34	AF152099	AF152099 Homo sapi
	40	56.8	3.5	22449	64	SPSNBCDE	Y11548 S.pristinae
	41	56.8	3.5	22449	64	SPSNBCGEN	X98690 S.pristinae
c	42	56.8	3.5	68848	2	MTV043	AL022004 Mycobacte
c	43	56.8	3.5	189579	48	AC023885	AC023885 Homo sapi
	44	56.8	3.5	245035	50	AC026782	AC026782 Homo sapi
	45	56.6	3.4	154868	35	AC011316	AC011316 Homo sapi

ALIGNMENTS

```
RESULT  1
AC018392/c
LOCUS       AC018392   140800 bp    DNA              HTG       09-SEP-2000
DEFINITION  Homo sapiens clone RP11-12K9, WORKING DRAFT SEQUENCE, 14 unordered
            pieces.
ACCESSION   AC018392
VERSION     AC018392.4  GI:10045345
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 140800)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-12K9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 140800)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
            Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W.,
            Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B.,
            Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L.,
            Karatas,A., Klein,J., Landers,T., Lehoczky,J., Lieu,C., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T.,
            O'Donnell,P., Peterson,K., Pierre,N., Pollara,V., Riley,R.,
            Rothman,D., Roy,A., Santos,R., Severy,P., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 9, 2000 this sequence version replaced gi:6649406.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3288
                Center clone name: 12_K_9
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 133050 bases at least Q40
                Consensus quality: 136489 bases at least Q30
                Consensus quality: 137937 bases at least Q20
                Insert size: 139000; agarose-fp
                Insert size: 139500; sum-of-contigs
                Quality coverage: 5.3 in Q20 bases; agarose-fp
                Quality coverage: 5.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     4056: contig of 4056 bp in length
            *    4057 4156: gap of      100 bp
            *    4157     5294: contig of 1138 bp in length
            *    5295 5394: gap of      100 bp
            *    5395     6816: contig of 1422 bp in length
            *    6817 6916: gap of      100 bp
            *    6917    57039: contig of 50123 bp in length
            *   57040 57139: gap of      100 bp
            *   57140    60223: contig of 3084 bp in length
            *   60224 60323: gap of      100 bp
            *   60324    63674: contig of 3351 bp in length
            *   63675 63774: gap of      100 bp
            *   63775    68200: contig of 4426 bp in length
            *   68201 68300: gap of      100 bp
            *   68301    76558: contig of 8258 bp in length
            *   76559 76658: gap of      100 bp
            *   76659    84493: contig of 7835 bp in length
            *   84494 84593: gap of      100 bp
            *   84594    94397: contig of 9804 bp in length
            *   94398 94497: gap of      100 bp
            *   94498   104204: contig of 9707 bp in length
            *  104205 104304: gap of      100 bp
            *  104305   118656: contig of 14352 bp in length
            *  118657 118756: gap of      100 bp
            *  118757   139302: contig of 20546 bp in length
            *  139303 139402: gap of      100 bp
            *  139403   140800: contig of 1398 bp in length.
FEATURES             Location/Qualifiers
     source          1. .140800
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-12K9"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .4056
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    4157. .5294
                     /note="assembly_fragment"
     misc_feature    5395. .6816
                     /note="assembly_fragment"
     misc_feature    6917. .57039
                     /note="assembly_fragment"
     misc_feature    57140. .60223
                     /note="assembly_fragment"
     misc_feature    60324. .63674
                     /note="assembly_fragment"
     misc_feature    63775. .68200
                     /note="assembly_fragment"
     misc_feature    68301. .76558
```

```
                     /note="assembly_fragment"
     misc_feature    76659. .84493
                     /note="assembly_fragment"
     misc_feature    84594. .94397
                     /note="assembly_fragment"
     misc_feature    94498. .104204
                     /note="assembly_fragment"
     misc_feature    104305. .118656
                     /note="assembly_fragment"
     misc_feature    118757. .139302
                     /note="assembly_fragment"
     misc_feature    139403. .140800
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
BASE COUNT    39565 a  29335 c  29337 g  41257 t   1306 others
ORIGIN
```

Query Match 89.7%; Score 1472.2; DB 38; Length 140800;
Best Local Similarity 97.8%; Pred. No. 6.2e-218;
Matches 1492; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy    109 ccgaccgccgcttccggccgcccaccaacctgcgcagcgtgtcgccctgggcctacagaa 168
          | | ||| ||    ||| |||| |   | || |        || ||   ||| ||||||
Db  22156 CAGCGCGCGGCGCGGGGCTGCCCCCAGGCGTGACCTCACCCGTGCTCTCTCCCTGCAGAA 22097

Qy    169 tctcctacgacccggcgaggtaccccaggtacctgcctgaagcctactgcctgtgccggg 228
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22096 TCTCCTACGACCCGGCGAGGTACCCCAGGTACCTGCCTGAAGCCTACTGCCTGTGCCGGG 22037

Qy    229 gctgcctgaccgggctgttcggcgaggaggacgtgcgcttccgcagcgcccctgtctaca 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  22036 GCTGCCTGACCGGGCTGTTCGGCGAGGAGGACGTGCGCTTCCGCAGCGCCCCTGTCTACA 21977

Qy    289 tgcccaccgtcgtcctgcgccgcacccccgcctgcgccggcggccgttccgtctacaccg 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21976 TGCCCACCGTCGTCCTGCGCCGCACCCCCGCCTGCGCCGGCGGCCGTTCCGTCTACACCG 21917

Qy    349 aggcctacgtcaccatccccgtgggctgcacctgcgtccccgagccggagaaggacgcag 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21916 AGGCCTACGTCACCATCCCCGTGGGCTGCACCTGCGTCCCCGAGCCGGAGAAGGACGCAG 21857

Qy    409 acagcatcaactccagcatcgacaaacagggcgccaagctcctgctgggccccaacgacg 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21856 ACAGCATCAACTCCAGCATCGACAAACAGGGCGCCAAGCTCCTGCTGGGCCCCAACGACG 21797

Qy    469 cgcccgctggcccctgaggccggtcctgccccgggaggtctccccggcccgcatcccgag 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21796 CGCCCGCTGGCCCCTGAGGCCGGTCCTGCCCCGGGAGGTCTCCCCGGCCCGCATCCCGAG 21737

Qy    529 gcgcccaagctggagccgcctggagggctcggtcggcgacctctgaagagagtgcaccga 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21736 GCGCCCAAGCTGGAGCCGCCTGGAGGGCTCGGTCGGCGACCTCTGAAGAGAGTGCACCGA 21677

Qy    589 gcaaaccaagtgccggagcaccagcgccgcctttccatggagactcgtaagcagcttcat 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21676 GCAAACCAAGTGCCGGAGCACCAGCGCCGCCTTTCCATGGAGACTCGTAAGCAGCTTCAT 21617

Qy    649 ctgacacgggcatccctggcttgcttttagctacaagcaagcagcgtggctggaagctga 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21616 CTGACACGGGCATCCCTGGCTTGCTTTTAGCTACAAGCAAGCAGCGTGGCTGGAAGCTGA 21557

Qy    709 tgggaaacgacccggcacgggcatcctgtgtgcggcccgcatggagggtttggaaaagtt 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21556 TGGGAAACGACCCGGCACGGGCATCCTGTGTGCGGCCCGCATGGAGGGTTTGGAAAAGTT 21497

Qy    769 cacggaggctccctgaggagcctctcagatcggctgctgcgggtgcagggcgtgactcac 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21496 CACGGAGGCTCCCTGAGGAGCCTCTCAGATCGGCTGCTGCGGGTGCAGGGCGTGACTCAC 21437

Qy    829 cgctgggtgcttgccaaagagatagggacgcatatgctttttaaagcaatctaaaaataa 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  21436 CGCTGGGTGCTTGCCAAAGAGATAGGGACGCATATGCTTTTTAAAGCAATCTAAAAATAA 21377

Qy    889 taataagtatagcgactatatacctacttttaaaatcaactgttttgaatagaggcagag 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21376 TAATAAGTATAGCGACTATATACCTACTTTTAAAATCAACTGTTTTGAATAGAGGCAGAG 21317

Qy    949 ctattttatattatcaaatgagagctactctgttacatttcttaacatataaacatcgtt 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21316 CTATTTTATATTATCAAATGAGAGCTACTCTGTTACATTTCTTAACATATAAACATCGTT 21257

Qy   1009 ttttacttcttctggtagaatttttttaaagcataattggaatccttggataaatttttgta 1068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21256 TTTTACTTCTTCTGGTAGAATTTTTTAAAGCATAATTGGAATCCTTGGATAAATTTTGTA 21197

Qy   1069 gctggtacactctggcctgggtctctgaattcagcctgtcaccgatggctgactgatgaa 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21196 GCTGGTACACTCTGGCCTGGGTCTCTGAATTCAGCCTGTCACCGATGGCTGACTGATGAA 21137

Qy   1129 atggacacgtctcatctgacccactcttccttccactgaaggtcttcacgggcctccagg 1188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21136 ATGGACACGTCTCATCTGACCCACTCTTCCTTCCACTGAAGGTCTTCACGGGCCTCCAGG 21077

Qy   1189 tggaccaaagggatgcacaggcggctcgcatgccccagggccagctaagagttccaaaga 1248
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21076 TGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGTTCCAAAGA 21017

Qy   1249 tctcagatttggttttagtcatgaatacataaacagtctcaaactcgcacaatttttttcc 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21016 TCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCGCACAATTTTTTCC 20957

Qy   1309 ccctttgaaagccactggggccaatttgtggttaagaggtggtgagataagaagtggaa 1368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  20956 CCCTTTTGAAAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGTGAGATAAGAAGTGGAA 20897

Qy   1369 cgtgacatctttgccagttgtcagaagaatccaagcaggtattggcttagttgtaagggc 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  20896 CGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCTTAGTTGTAAGGGC 20837

Qy   1429 tttaggatcaggctgaatatgaggacaaagtgggccacgttagcatctgcagagatcaat 1488
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  20836 TTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATCTGCAGAGATCAAT 20777

Qy   1489 ctggaggcttctgtttctgcattctgccacgagagctaggtccttgatcttttctttaga 1548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  20776 CTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGATCTTTTCTTTAGA 20717

Qy   1549 ttgaaagtctgtctctgaacacaattatttgtaaaagttagtagttctttttttaaatcat 1608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  20716 TTGAAAGTCTGTCTCTGAACACAATTATTTGTAAAAGTTAGTAGTTCTTTTTTAAATCAT 20657

Qy   1609 taaaagaggcttgctgaaaaaaaaa 1633
          |||||||||||||||||  | | |
Db  20656 TAAAAGAGGCTTGCTGAAGGATACA 20632
```

```
RESULT   2
AL161772/c
LOCUS       AL161772   142376 bp    DNA             HTG       23-AUG-2000
DEFINITION  Homo sapiens chromosome 13 clone RP11-172H24, *** SEQUENCING IN
            PROGRESS ***, 19 unordered pieces.
ACCESSION   AL161772
VERSION     AL161772.5  GI:9796851
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 142376)
  AUTHORS   Burton,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
```

```
COMMENT     On Aug 12, 2000 this sequence version replaced gi:9212853.
            ---------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ---------- Project Information
            Center project name: bA172H24
            ---------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 129659 bases at least Q40
            Consensus quality: 134544 bases at least Q30
            Consensus quality: 137507 bases at least Q20
            Insert size: 140576; sum-of-contigs
            Insert size: 159457; 9.8% error; agarose-fp
            Quality coverage: 3.18x in Q20 bases; sum-of-contigs Quality
            coverage: 2.80x in Q20 bases; agarose-fp
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 19 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     7156: contig of 7156 bp in length
            *    7157 7256: gap of      100 bp
            *    7257    11058: contig of 3802 bp in length
            *   11059 11158: gap of      100 bp
            *   11159    22535: contig of 11377 bp in length
            *   22536 22635: gap of      100 bp
            *   22636    27293: contig of 4658 bp in length
            *   27294 27393: gap of      100 bp
            *   27394    33036: contig of 5643 bp in length
            *   33037 33136: gap of      100 bp
            *   33137    44104: contig of 10968 bp in length
            *   44105 44204: gap of      100 bp
            *   44205    51504: contig of 7300 bp in length
            *   51505 51604: gap of      100 bp
            *   51605    53871: contig of 2267 bp in length
            *   53872 53971: gap of      100 bp
            *   53972    60840: contig of 6869 bp in length
            *   60841 60940: gap of      100 bp
            *   60941    64255: contig of 3315 bp in length
            *   64256 64355: gap of      100 bp
            *   64356    69706: contig of 5351 bp in length
            *   69707 69806: gap of      100 bp
            *   69807    77391: contig of 7585 bp in length
            *   77392 77491: gap of      100 bp
            *   77492    80766: contig of 3275 bp in length
            *   80767 80866: gap of      100 bp
            *   80867    91820: contig of 10954 bp in length
            *   91821 91920: gap of      100 bp
            *   91921    94596: contig of 2676 bp in length
            *   94597 94696: gap of      100 bp
            *   94697   102331: contig of 7635 bp in length
            *  102332 102431: gap of      100 bp
            *  102432   114250: contig of 11819 bp in length
            *  114251 114350: gap of      100 bp
            *  114351   133624: contig of 19274 bp in length
            *  133625 133724: gap of      100 bp
            *  133725   142376: contig of 8652 bp in length.
FEATURES             Location/Qualifiers
     source          1. .142376
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="RP11-172H24"
                     /clone_lib="RPCI-11.1"
     misc_feature    1. .7156
                     /note="assembly_fragment:00756
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    7257. .11058
                     /note="assembly_fragment:01021
                     fragment_chain:1"
     misc_feature    11159. .22535
                     /note="assembly_fragment:00324
                     fragment_chain:1"
     misc_feature    22636. .27293
                     /note="assembly_fragment:00287
                     fragment_chain:2"
     misc_feature    27394. .33036
                     /note="assembly_fragment:00179
                     fragment_chain:2"
     misc_feature    33137. .44104
                     /note="assembly_fragment:00048
                     fragment_chain:2"
     misc_feature    44205. .51504
                     /note="assembly_fragment:00045
                     fragment_chain:3"
     misc_feature    51605. .53871
                     /note="assembly_fragment:00128
                     fragment_chain:3"
     misc_feature    53972. .60840
                     /note="assembly_fragment:00702
                     fragment_chain:4"
     misc_feature    60941. .64255
                     /note="assembly_fragment:00935
                     fragment_chain:4"
     misc_feature    64356. .69706
                     /note="assembly_fragment:00869
                     fragment_chain:5"
     misc_feature    69807. .77391
                     /note="assembly_fragment:00527
                     fragment_chain:5"
     misc_feature    77492. .80766
                     /note="assembly_fragment:00239"
     misc_feature    80867. .91820
                     /note="assembly_fragment:00340"
     misc_feature    91921. .94596
                     /note="assembly_fragment:00427"
     misc_feature    94697. .102331
                     /note="assembly_fragment:00554"
     misc_feature    102432. .114250
                     /note="assembly_fragment:00637"
     misc_feature    114351. .133624
                     /note="assembly_fragment:01082"
     misc_feature    133725. .142376
                     /note="assembly_fragment:01087"
BASE COUNT    41636 a  28499 c  28976 g  41458 t   1807 others
ORIGIN


  Query Match             88.6%;  Score 1455.2;  DB 70;  Length 142376;
  Best Local Similarity   97.1%;  Pred. No. 2.6e-215;
  Matches 1481;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy     109 ccgaccgccgcttccggccgcccaccaacctgcgcagcgtgtcgccctgggcctacagaa 168
           | |  ||| ||    ||| |||| |   | ||  |         || ||   ||| |||||
Db  131140 CAGCGCGCGGCGCGGGGCTGCCCCCAGGCGTGACCTCACCCGTGCTCTCTCCCTGCAGAA 131081

Qy     169 tctcctacgacccggcgaggtaccccaggtacctgcctgaagcctactgcctgtgccggg 228
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  131080 TCTCCTACGACCCGGCGAGGTACCCCAGGTACCTGCCTGAAGCCTACTGCCTGTGCCGGG 131021

Qy     229 gctgcctgaccgggctgttcggcgaggaggacgtgcgcttccgcagcgcccctgtctaca 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  131020 GCTGCCTGACCGGGCTGTTCGGCGAGGAGGACGTGCGCTTCCGCAGCGCCCCTGTCTACA 130961

Qy     289 tgcccaccgtcgtcctgcgccgcaccccgcctgcgccggcggccgttccgtctacaccg 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 130960 TGCCCACCGTCGTCCTGCGCCGCACCCCCGCCTGCGCCGGCGGCCGTTCCGTCTACACCG 130901

Qy    349 aggcctacgtcaccatccccgtgggctgcacctgcgtccccgagccggagaaggacgcag 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130900 AGGCCTACGTCACCATCCCCGTGGGCTGCACCTGCGTCCCCGAGCCGGAGAAGGACGCAG 130841

Qy    409 acagcatcaactccagcatcgacaaacagggcgccaagctcctgctgggccccaacgacg 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130840 ACAGCATCAACTCCAGCATCGACAAACAGGGCGCCAAGCTCCTGCTGGGCCCCAACGACG 130781

Qy    469 cgcccgctggcccctgaggccggtcctgccccgggaggtctccccggcccgcatcccgag 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130780 CGCCCGCTGGCCCCTGAGGCCGGTCCTGCCCCGGGAGGTCTCCCCGGCCCGCATCCCGAG 130721

Qy    529 gcgcccaagctggagccgcctggagggctcggtcggcgacctctgaagagagtgcaccga 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130720 GCGCCCAAGCTGGAGCCGCCTGGAGGGCTCGGTCGGCGACCTCTGAAGAGAGTGCACCGA 130661

Qy    589 gcaaaccaagtgccggagcaccagcgccgcctttccatggagactcgtaagcagcttcat 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130660 GCAAACCAAGTGCCGGAGCACCAGCGCCGCCTTTCCATGGAGACTCGTAAGCAGCTTCAT 130601

Qy    649 ctgacacgggcatccctggcttgcttttagctacaagcaagcagcgtggctggaagctga 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130600 CTGACACGGGCATCCCTGGCTTGCTTTTAGCTACAAGCAAGCAGCGTGGCTGGAAGCTGA 130541

Qy    709 tgggaaacgacccggcacgggcatcctgtgtgcggcccgcatggagggtttggaaaagtt 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130540 TGGGAAACGACCCGGCACGGGCATCCTGTGTGCGGCCCGCATGGAGGGTTTGGAAAAGTT 130481

Qy    769 cacggaggctccctgaggagcctctcagatcggctgctgcgggtgcagggcgtgactcac 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130480 CACGGAGGCTCCCTGAGGAGCCTCTCAGATCGGCTGCTGCGGGTGCAGGGCGTGACTCAC 130421

Qy    829 cgctgggtgcttgccaaagagatagggacgcatatgctttttaaagcaatctaaaaataa 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130420 CGCTGGGTGCTTGCCAAAGAGATAGGGACGCATATGCTTTTTAAAGCAATCTAAAAATAA 130361

Qy    889 taataagtatagcgactatatacctacttttaaaatcaactgttttgaatagaggcagag 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130360 TAATAAGTATAGCGACTATATACCTACTTTTAAAATCAACTGTTTTGAATAGAGGCAGAG 130301

Qy    949 ctattttatattatcaaatgagagctactctgttacatttcttaacatataaacatcgtt 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130300 CTATTTTATATTATCAAATGAGAGCTACTCTGTTACATTTCTTAACATATAAACATCGTT 130241

Qy   1009 ttttacttcttctggtagaatttttttaaagcataattggaatccttggataaattttgta 1068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130240 TTTTACTTCTTCTGGTAGAATTTTTTAAAGCATAATTGGAATCCTTGGATAAATTTTGTA 130181

Qy   1069 gctggtacactctggcctgggtctctgaattcagcctgtcaccgatggctgactgatgaa 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130180 GCTGGTACACTCTGGCCTGGGTCTCTGAATTCAGCCTGTCACCGATGGCTGACTGATGAA 130121

Qy   1129 atggacacgtctcatctgacccactcttccttccactgaaggtcttcacgggcctccagg 1188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130120 ATGGACACGTCTCATCTGACCCACTCTTCCTTCCACTGAAGGTCTTCACGGGCCTCCAGG 130061

Qy   1189 tggaccaaagggatgcacaggcggctcgcatgccccagggccagctaagagttccaaaga 1248
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130060 TGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGTTCCAAAGA 130001

Qy   1249 tctcagatttggttttagtcatgaatacataaacagtctcaaactcgcacaatttttttcc 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130000 TCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCGCACAATTTTTTCC 129941

Qy   1309 ccctttttgaaagccactggggccaatttgtggttaagaggtggtgagataagaagtggaa 1368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129940 CCCTTTTGAAAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGTGAGATAAGAAGTGGAA 129881

Qy   1369 cgtgacatctttgccagttgtcagaagaatccaagcaggtattggcttagttgtaagggc 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129880 CGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCTTAGTTGTAAGGGC 129821

Qy   1429 tttaggatcaggctgaatatgaggacaaagtgggccacgttagcatctgcagagatcaat 1488
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129820 TTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATCTGCAGAGATCAAT 129761

Qy   1489 ctggaggcttctgtttctgcattctgccacgagagctaggtccttgatcttttctttaga 1548
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 129760 CTGAAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGATCTTTTCTTTAGA 129701

Qy   1549 ttgaaagtctgtctctgaacacaattatttgtaaaagttagtagttctttttaaatcat 1608
          |||||||||||| ||      ||||||| |||||||||||||||  ||||||||||||
Db 129700 TTGAAAGTCTGTTTCGATCTNCAATTATATGTAAAAGTTAGTAGTCTTTTTTTAAATCAT 129641

Qy   1609 taaaagaggcttgctgaaaaaaaaa 1633
          |||||||||||||||||  | | |
Db 129640 TAAAAGAGGCTTGCTGAAGGATACA 129616
```

(5)

```
RESULT   3
HUMNOTIB/c
LOCUS       HUMNOTIB      629 bp    mRNA            PRI       13-DEC-1993
DEFINITION  Human chromosome-specific mRNA.
ACCESSION   L23207
VERSION     L23207.1  GI:434048
KEYWORDS    .
SOURCE      Homo sapiens female 3 month post natal entire brain cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 629)
  AUTHORS   Bonaldo,M., Soares,M.-B. and Warburton,D.
  TITLE     Selection of chromosome-specific cDNAs and their corresponding
            genomic clones
  JOURNAL   Unpublished (1993)
COMMENT     Chromosome 13q11.
FEATURES             Location/Qualifiers
     source          1. .629
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="3 month post natal"
                     /sex="female"
                     /tissue_type="entire brain"
BASE COUNT      189 a    149 c    128 g    163 t
ORIGIN
```

Query Match 38.0%; Score 624.4; DB 85; Length 629;
Best Local Similarity 99.8%; Pred. No. 5.6e-87;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy   1001 acatcgttttttacttcttctggtagaatttttaaagcataattggaatccttggataa 1060
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    629 ACATCGTTTTTTACTTCTTCTGGTAGAATTTGTTAAAGCATAATTGGAATCCTTGGATAA 570

Qy   1061 attttgtagctggtacactctggcctgggtctctgaattcagcctgtcaccgatggctga 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     569 ATTTTGTAGCTGGTACACTCTGGCCTGGGTCTCTGAATTCAGCCTGTCACCGATGGCTGA 510

Qy   1121 ctgatgaaatggacacgtctcatctgacccactcttccttccactgaaggtcttcacggg 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     509 CTGATGAAATGGACACGTCTCATCTGACCCACTCTTCCTTCCACTGAAGGTCTTCACGGG 450

Qy   1181 cctccaggtggaccaaagggatgcacaggcggctcgcatgccccagggccagctaagagt 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     449 CCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGT 390

Qy   1241 tccaaagatctcagatttggttttagtcatgaatacataaacagtctcaaactcgcacaa 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     389 TCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCGCACAA 330

Qy   1301 ttttttcccccttttgaaagccactggggccaatttgtggttaagaggtggtgagataag 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     329 TTTTTTCCCCCTTTTGAAAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGTGAGATAAG 270
```

```
Qy   1361 aagtggaacgtgacatctttgccagttgtcagaagaatccaagcaggtattggcttagtt 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    269 AAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCTTAGTT 210

Qy   1421 gtaagggctttaggatcaggctgaatatgaggacaaagtgggccacgttagcatctgcag 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    209 GTAAGGGCTTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATCTGCAG 150

Qy   1481 agatcaatctggaggcttctgtttctgcattctgccacgagagctaggtccttgatcttt 1540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149 AGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGATCTTT 90

Qy   1541 tctttagattgaaagtctgtctctgaacacaattatttgtaaaagttagtagttctttttt 1600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     89 TCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATTTGTAAAAGTTAGTAGTTCTTTTT 30

Qy   1601 taaatcattaaaagaggcttgctgaa 1626
          |||||||||||||||||||||||||,
Db     29 TAAATCATTAAAAGAGGCTTGCTGAA 4
```

RESULT 4
HUMCH13C/c
LOCUS HUMCH13C 296 bp mRNA PRI 24-MAR-1994
DEFINITION Homo sapiens (clone ch13cDNA19B) mRNA sequence.
ACCESSION L30108
VERSION L30108.1 GI:463092
KEYWORDS .
SOURCE Homo sapiens female 3 month post-natal brain cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 296)
AUTHORS Bonaldo,M., Soares,M.-B. and Warburton,D.
TITLE Selection of chromosome-specific cDNAs and their corresponding genomic clones
JOURNAL Unpublished (1993)
FEATURES Location/Qualifiers
source 1. .296
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="3 month post-natal"
/sex="female"
/tissue_type="brain"
/map="13q11"
misc_feature 1. .296
BASE COUNT 84 a 66 c 73 g 73 t
ORIGIN

Query Match 17.9%; Score 294.4; DB 85; Length 296;
Best Local Similarity 99.7%; Pred. No. 4e-36;
Matches 295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy   1001 acatcgtttttttacttcttctgctagaatttttttaaagcataattggaatccttggataa 1060
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    296 ACATCGTTTTTTACTTCTTCTGGTAGAATTTGTTAAAGCATAATTGGAATCCTTGGATAA 237

Qy   1061 attttgtagctggtacactctggcctgggtctctgaattcagcctgtcaccgatggctga 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    236 ATTTTGTAGCTGGTACACTCTGGCCTGGGTCTCTGAATTCAGCCTGTCACCGATGGCTGA 177

Qy   1121 ctgatgaaatggacacgtctcatctgacccactcttccttccactgaaggtcttcacggg 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    176 CTGATGAAATGGACACGTCTCATCTGACCCACTCTTCCTTCCACTGAAGGTCTTCACGGG 117

Qy   1181 cctccaggtggaccaaagggatgcacaggcggctcgcatgccccagggccagctaagagt 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    116 CCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGT 57

Qy   1241 tccaaagatctcagatttggtttagtcatgaatacataaacagtctcaaactcgc 1296
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     56 TCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCGC 1
```

(6)

RESULT 5
HUMNOTIA/c
LOCUS HUMNOTIA 323 bp mRNA PRI 13-DEC-1993
DEFINITION Human chromosome specific mRNA.
ACCESSION L23206
VERSION L23206.1 GI:434047
KEYWORDS .
SOURCE Homo sapiens female 3 month post natal entire brain cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 323)
AUTHORS Bonaldo,M., Soares,M.-B. and Warburton,D.
TITLE Selection of chromosome-specific cDNAs and their corresponding genomic clones
JOURNAL Unpublished (1993)
COMMENT Chromosome 13q11.
FEATURES Location/Qualifiers
source 1. .323
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="3 month post natal"
/sex="female"
/tissue_type="entire brain"
BASE COUNT 41 a 120 c 119 g 43 t
ORIGIN

Query Match 15.4%; Score 252.8; DB 85; Length 323;
Best Local Similarity 99.2%; Pred. No. 9.9e-30;
Matches 254; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy     21 gccgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 80
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    256 GGCGTGCTCAGTGCCTTCCACCACACGCTGCAGCTGGGGCCGCGTGAGCAGGCGCGCAAC 197

Qy     81 gcgagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 140
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    196 GCGAGCTGCCCGGCAGGGGGCAGGCCCGCCGACCGCCGCTTCCGGACGCCCACCAACCTG 137

Qy    141 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 CGCAGCGTGTCGCCCTGGGCCTACAGAATCTCCTACGACCCGGCGAGGTACCCCAGGTAC 77

Qy    201 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 CTGCCTGAAGCCTACTGCCTGTGCCGGGGCTGCCTGACCGGGCTGTTCGGCGAGGAGGAC 17

Qy    261 gtgcgcttccgcagcg 276
          ||||||||||||||||
Db     16 GTGCGCTTCCGCAGCG 1
```

(7)

RESULT 6
G31083/c
LOCUS G31083 250 bp DNA STS 31-DEC-1996
DEFINITION human STS SHGC-19175, sequence tagged site.
ACCESSION G31083
VERSION G31083.1 GI:1758484
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Myers,R.M.
JOURNAL Unpublished (1996)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)

Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: CAAGGACCTAGCTCTCGTGG
Primer B: GAAGAATCCAAGCAGGTATTGG
STS size: 143
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs derived from H29012--Merck/UniEST.

```
FEATURES             Location/Qualifiers
     source          1. .250
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13"
     STS             108. .250
     primer_bind     108. .127
     primer_bind     complement(229. .250)
BASE COUNT       77 a     58 c     38 g     77 t
ORIGIN
```

```
  Query Match             14.9%;  Score 245.2;  DB 75;  Length 250;
  Best Local Similarity   98.8%;  Pred. No. 1.5e-28;
  Matches  247;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1392 gaagaatccaagcaggtattggcttagttgtaagggctttaggatcaggctgaatatgag 1451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    250 GAAGAATCCAAGCAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATCAGGCTGAATATGAG 191

Qy   1452 gacaaagtgggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcatt 1511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    190 GACAAAGTGGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCATT 131

Qy   1512 ctgccacgagagctaggtccttgatcttttctttagattgaaagtctgtctctgaacaca 1571
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 CTGCCACGAGAGCTAGGTCCTTGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACACA 71

Qy   1572 attatttgtaaaagttagtagttctttttttaaatcattaaaagaggcttgctgaaaaaaa 1631
          |||||||||||||||||||||||||||||||||||||||||||||||||||||  | |
Db     70 ATTATTTGTAAAAGTTAGTAGTTCTTTTTTAAATCATTAAAAGAGGCTTGCTGAAGGATA 11

Qy   1632 aaaaaaaaaa 1641
          ||||||||||
Db     10 AAAAAAAAAA 1
```

```
RESULT   7
AL356002/c
LOCUS       AL356002   162005 bp    DNA             HTG       10-AUG-2000
DEFINITION  Homo sapiens chromosome 1 clone RP11-324K19, *** SEQUENCING IN
            PROGRESS ***, 18 unordered pieces.
ACCESSION   AL356002
VERSION     AL356002.4  GI:9797411
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 162005)
  AUTHORS   Plumb,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 12, 2000 this sequence version replaced gi:9213703.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA324K19
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 150712 bases at least Q40
            Consensus quality: 155209 bases at least Q30
            Consensus quality: 157665 bases at least Q20
            Insert size: 160305; sum-of-contigs
            Insert size: 170079; 1.7% error; agarose-fp
            Quality coverage: 3.29x in Q20 bases; sum-of-contigs Quality
            coverage: 3.10x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     4977: contig of 4977 bp in length
            *    4978 5077: gap of      100 bp
            *    5078    14284: contig of 9207 bp in length
            *   14285 14384: gap of      100 bp
            *   14385    20588: contig of 6204 bp in length
            *   20589 20688: gap of      100 bp
            *   20689    27191: contig of 6503 bp in length
            *   27192 27291: gap of      100 bp
            *   27292    39205: contig of 11914 bp in length
            *   39206 39305: gap of      100 bp
            *   39306    43223: contig of 3918 bp in length
            *   43224 43323: gap of      100 bp
            *   43324    46284: contig of 2961 bp in length
            *   46285 46384: gap of      100 bp
            *   46385    73430: contig of 27046 bp in length
            *   73431 73530: gap of      100 bp
            *   73531   104566: contig of 31036 bp in length
            *  104567 104666: gap of      100 bp
            *  104667   109641: contig of 4975 bp in length
            *  109642 109741: gap of      100 bp
            *  109742   113089: contig of 3348 bp in length
            *  113090 113189: gap of      100 bp
            *  113190   132251: contig of 19062 bp in length
            *  132252 132351: gap of      100 bp
            *  132352   138017: contig of 5666 bp in length
            *  138018 138117: gap of      100 bp
            *  138118   141454: contig of 3337 bp in length
            *  141455 141554: gap of      100 bp
            *  141555   145269: contig of 3715 bp in length
            *  145270 145369: gap of      100 bp
            *  145370   150364: contig of 4995 bp in length
            *  150365 150464: gap of      100 bp
```

```
          *   150465   153914: contig of 3450 bp in length
          *   153915 154014: gap of      100 bp
          *   154015   162005: contig of 7991 bp in length.
FEATURES             Location/Qualifiers
     source          1. .162005
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-324K19"
                     /clone_lib="RPCI-11.2"
     misc_feature    1. .4977
                     /note="assembly_fragment:00867
                     clone_end:SP6
                     vector_side:left"
     misc_feature    5078. .14284
                     /note="assembly_fragment:00646
                     fragment_chain:1"
     misc_feature    14385. .20588
                     /note="assembly_fragment:00037
                     fragment_chain:1"
     misc_feature    20689. .27191
                     /note="assembly_fragment:00574
                     fragment_chain:1"
     misc_feature    27292. .39205
                     /note="assembly_fragment:00176
                     fragment_chain:2"
     misc_feature    39306. .43223
                     /note="assembly_fragment:00772
                     fragment_chain:2"
     misc_feature    43324. .46284
                     /note="assembly_fragment:00453
                     fragment_chain:3"
     misc_feature    46385. .73430
                     /note="assembly_fragment:00029
                     fragment_chain:3"
     misc_feature    73531. .104566
                     /note="assembly_fragment:00717
                     fragment_chain:4"
     misc_feature    104667. .109641
                     /note="assembly_fragment:00524
                     fragment_chain:4"
     misc_feature    109742. .113089
                     /note="assembly_fragment:00101"
     misc_feature    113190. .132251
                     /note="assembly_fragment:00160"
     misc_feature    132352. .138017
                     /note="assembly_fragment:00225"
     misc_feature    138118. .141454
                     /note="assembly_fragment:00400"
     misc_feature    141555. .145269
                     /note="assembly_fragment:00418"
     misc_feature    145370. .150364
                     /note="assembly_fragment:01084"
     misc_feature    150465. .153914
                     /note="assembly_fragment:00162
                     fragment_chain:5"
     misc_feature    154015. .162005
                     /note="assembly_fragment:00655
                     fragment_chain:5
                .    clone_end:T7
                     vector_side:right"
BASE COUNT    49657 a  33131 c  32596 g  44854 t   1767 others
ORIGIN


  Query Match             4.7%;  Score 77;  DB 71;  Length 162005;
  Best Local Similarity   43.1%;  Pred. No. 0.0033;
  Matches  233;  Conservative    0;  Mismatches  308;  Indels    0;  Gaps    0;

Qy      18 cgtgccgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgc 77
           | | ||   |||  | ||| ||  |   |    |          |         | |
Db 151191 CTTCCCTCCCTCCCTNCCTNCCTNCCNNCCNNNNCCCCNNCCNNCCCNNCNCCCNN 151132
```

```
Qy      78 aacgcgagctgcccggcagggggcaggcccgccgaccgccgcttccggccgcccaccaac 137
           | |  |  |  |||   |   | | |||| |   ||| |  |||   |||     |
Db 151131 NCCNCCGGGGGGGCGGGCNGCCCCGGCCCCGGCCCGGGCCCCCGCCGCGGCCCCGGGCCC 151072

Qy     138 ctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccagg 197
           | | |  ||    |  || ||| |  | |       |  || || |||  ||  || | |
Db 151071 CGGGGGCGCCCCCCCNCCCGGGGCGCCCGCGGGCGCGNCGCCCGGGCCGGGGGCCGCGGC 151012

Qy     198 tacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggag 257
           |   || |  | |  |  | | |||||||  |     |      |   |||| | |
Db 151011 CNCGCCCCNGCCGGCCGCGCGCGGGGCCGGGGNNGGGGCGCGCCCGCGNCGGGCGCGCCG 150952

Qy     258 gacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccc 317
           | |    ||  ||   ||    | | |  |   |||  || |      |  | | |||||
Db 150951 GCCCNNGGCCCCCNGGGCCGGGCGGCCCGCCNCCCCGNCGNCCCGGGCCCGCCCCCCCCC 150892

Qy     318 gcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgc 377
           | | | ||| ||| ||    || |  |   | || |  |  | ||    |||  ||   |
Db 150891 GGCGGGGCCCGCGNCCCCGGCGCCCCCCGGGGGGGCGGCCCCCCCCGNNCCGCCGGGNCC 150832

Qy     378 acctgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaaacag 437
           |    | ||  ||  | |      || | | |        ||  |  ||     | |
Db 150831 CGGCGGNCCGCCCGGCGCGGGGGNNGGGCGGGCGGGCGGGNNCCCGCCCGCGCGCGCCGG 150772

Qy     438 ggcgccaagctcctgctgggccccaacgacgcgcccgctggcccctgaggccggtcctgc 497
           ||  ||      || ||  ||||||  |  | ||||      |||  |    ||  || ||
Db 150771 GGGCCCGCCGGCCGGCCCGGCCCCCGCCCCCCGCCGCGCCCCCCGGGCCCGCGCCCCCGC 150712

Qy     498 cccgggaggtctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggct 557
           |  ||     | |||| |||| |  |||  ||  ||   ||  |   | |   | || |
Db 150711 CGGGGNCCCCCCCCCCCGCCCCCGCCCCCGGGNNCCNCGGCGCGCCGCCCGCCGGGGCCC 150652

Qy     558 c 558
           |
Db 150651 C 150651


RESULT   8
AF152098
LOCUS       AF152098      687 bp    mRNA            PRI       27-JAN-2000
DEFINITION  Homo sapiens interleukin 17B mRNA, complete cds.
ACCESSION   AF152098
VERSION     AF152098.1  GI:6759912
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   Li,H., Chen,J., Huang,A., Stinson,J., Heldens,S., Foster,J.,
            Dowd,P., Gurney,A.L. and Wood,W.I.
  TITLE     Cloning and characterization of IL-17B and IL-17C, two new members
            of the IL-17 cytokine family
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (2), 773-778 (2000)
  MEDLINE   20105548
REFERENCE   2  (bases 1 to 687)
  AUTHORS   Li,H., Chen,J., Stinson,J. and Wood,W.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-1999) Molecular Biology, Genentech, Inc., 1 DNA
            Way, S. San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .687
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q32-q34"
                     /clone="DNA59294"
     CDS             42. .584
                     /note="IL-17B"
                     /codon_start=1
                     /product="interleukin 17B"
                     /protein_id="AAF28104.1"
                     /db_xref="GI:6759913"
```

```
                     /translation="MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQ
                     VPLDLVSRMKPYARMEEYERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWG
                     YSINHDPSRIPVDLPEARCLCLGCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRT
                     GPCRQRAVMETIAVGCTCIF"
BASE COUNT      142 a    216 c    209 g    120 t
ORIGIN


  Query Match             4.5%;  Score 73.2;  DB 34;  Length 687;
  Best Local Similarity   58.1%;  Pred. No. 0.038;
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy     133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
           |||||  | | ||| |||| ||||||| |||||| |||  | |||||||     |   ||
Db     319 CCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCCGTATCC 378

Qy     193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
           ||  | |||||||  || ||    ||||||| | ||||||  ||| |    |   |
Db     379 CCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACCATGC 438

Qy     253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
           ||||||||    || |    ||||  || || | ||      | ||      |  ||  |
Db     439 AGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCTGTGCGCCGCCGCC 498

Qy     307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
           | || | ||  |   |  |   || | || |  | |   ||  | |   |  ||||||
Db     499 TCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG 558

Qy     367 ccgtgggctgcacctgcgtc 386
           | ||||||||||||||| ||
Db     559 CTGTGGGCTGCACCTGCATC 578


RESULT  9
AF184969
LOCUS       AF184969      688 bp    mRNA            PRI       13-JAN-2000
DEFINITION  Homo sapiens cytokine-like protein ZCYTO7 (ZCYTO7) mRNA, complete
            cds.
ACCESSION   AF184969
VERSION     AF184969.1  GI:6013320
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 688)
  AUTHORS   Presnell,S., Gilbert,T., Whitmore,T., Foster,D., Hart,C.,
            Lehner,J., Martinez,T., Hoffman,R. and O'Hara,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-1999) Biomolecular Informatics, ZymoGenetics,
            Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA
FEATURES             Location/Qualifiers
     source          1. .688
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /db_xref="dbEST:W74558"
                     /map="5q31.3-q32"
                     /clone="IMAGE:344649"
     gene            1. .688
                     /gene="ZCYTO7"
     CDS             42. .584
                     /gene="ZCYTO7"
                     /codon_start=1
                     /product="cytokine-like protein ZCYTO7"
                     /protein_id="AAF01318.1"
                     /db_xref="GI:6013321"
                     /translation="MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQ
                     VPLDLVSRMKPYARMEEYERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWG
                     YSINHDPSRIPVDLPEARCLCLGCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRT
                     GPCRQRAVMETIAVGCTCIF"
BASE COUNT      142 a    217 c    209 g    120 t
ORIGIN
```

```
  Query Match             4.5%;  Score 73.2;  DB 34;  Length 688;
  Best Local Similarity   58.1%;  Pred. No. 0.038;
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy     133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
           |||||  | | ||| |||| ||||||| |||||| |||  | |||||||     |   ||
Db     319 CCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCCGTATCC 378

Qy     193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
           ||  | |||||||  || ||    ||||||| | ||||||  ||| |    |   |
Db     379 CCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACCATGC 438

Qy     253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
           ||||||||    || |    ||||  || || | ||      | ||      |  ||  |
Db     439 AGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCTGTGCGCCGCCGCC 498

Qy     307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
           | || | ||  |   |  |   || | || |  | |   ||  | |   |  ||||||
Db     499 TCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG 558

Qy     367 ccgtgggctgcacctgcgtc 386
           | ||||||||||||||| ||
Db     559 CTGTGGGCTGCACCTGCATC 578


RESULT  10
AF212311
LOCUS       AF212311      711 bp    mRNA            PRI       26-JUN-2000
DEFINITION  Homo sapiens interleukin 20 (IL20) mRNA, complete cds.
ACCESSION   AF212311
VERSION     AF212311.1  GI:8705219
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Shi,Y., Ullrich,S.J., Zhang,J., Connolly,K., Grzegorzewski,K.J.,
            Barber,M.C., Wang,W., Wathen,K., Hodge,V., Fisher,C.L., Olsen,H.,
            Ruben,S.M., Knyazev,I., Cho,Y.H., Kao,V., Wilkinson,K.A.,
            Carrell,J.A. and Ebner,R.
  TITLE     A novel cytokine receptor-ligand pair. IDENTIFICATION, MOLECULAR
            CHARACTERIZATION, AND IN VIVO IMMUNOMODULATORY ACTIVITY
  JOURNAL   J. Biol. Chem. 275 (25), 19167-19176 (2000)
  MEDLINE   20317118
REFERENCE   2  (bases 1 to 711)
  AUTHORS   Shi,Y., Zhang,J., Ullrich,S.J., Connolly,K., Grzegorzewski,K.J.,
            Wang,W., Wathen,K., Hodge,V., Barber,M.C., Buergin,M., Kao,V.,
            Wilkinson,K.A., Carrell,J.A., Ruben,S.M. and and Ebner,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-DEC-1999) Molecular Biology, Human Genome Sciences,
            Inc., 9410 Key West Avenue, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.2"
     gene            1. .711
                     /gene="IL20"
     CDS             51. .593
                     /gene="IL20"
                     /codon_start=1
                     /product="interleukin 20"
                     /protein_id="AAF78775.1"
                     /db_xref="GI:8705220"
                     /translation="MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQ
                     VPLDLVSRMKPYARMEEYERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWG
                     YSINHDPSRIPVDLPEARCLCLGCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRT
                     GPCRQRAVMETIAVGCTCIF"
BASE COUNT      158 a    217 c    212 g    124 t
ORIGIN
```

```
  Query Match             4.5%;  Score 73.2;  DB 34;  Length 711;
  Best Local Similarity   58.1%;  Pred. No. 0.037;
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy     133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
           |||||  | | ||| |||| ||||||| |||||| |||  | |||||||     |   ||
Db     328 CCAACAAGAGGAGCCTGTCTCCGTGGGGCTACAGCATCAACCACGACCCCAGCCGTATCC 387

Qy     193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
           ||  | |||||||| || ||    |||||||| | ||||||  ||| |     |   |
Db     388 CCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACCATGC 447

Qy     253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
           ||||||||    || |    ||||  || || | ||      | ||      |   || |
Db     448 AGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCTGTGCGCCGCCGCC 507

Qy     307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
             | || | ||  |   | |   || | || |  | |   ||  | |   |  ||||||
Db     508 TCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG 567

Qy     367 ccgtgggctgcacctgcgtc 386
           | ||||||||||||||| ||
Db     568 CTGTGGGCTGCACCTGCATC 587
```

```
RESULT  11
SC7H2/c
LOCUS       SC7H2       42655 bp    DNA             BCT       02-AUG-1999
DEFINITION  Streptomyces coelicolor cosmid 7H2.
ACCESSION   AL109732
VERSION     AL109732.1  GI:5689879
KEYWORDS    amino acid decarboxylase; aminopeptidase; ATP-binding RNA helicase;
            formate dehydrogenase; integral membrane; monooxygenase; mutase;
            periplasmic branched chain amino acid binding; permease; polA, DNA
            polymerase I; pyk, pyruvate kinase; response regulator; rpsA, 30S
            ribosomal protein S1; secreted nucleotidase.
SOURCE      Streptomyces coelicolor A3(2).
  ORGANISM  Streptomyces coelicolor A3(2)
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 42655)
  AUTHORS   Redenbach,M., Kieser,H.M., Denapaite,D., Eichner,A., Cullum,J.,
            Kinashi,H. and Hopwood,D.A.
  TITLE     A set of ordered cosmids and a detailed genetic and physical map
            for the 8 Mb Streptomyces coelicolor A3(2) chromosome
  JOURNAL   Mol. Microbiol. 21 (1), 77-96 (1996)
  MEDLINE   97000351
REFERENCE   2  (bases 1 to 42655)
  AUTHORS   Saunders,D.C. and Harris,D.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 42655)
  AUTHORS   Bentley,S.D., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-1999) Streptomyces coelicolor sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
            CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
            David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
            Colney, Norwich, Norfolk NR4 7UH, UK
COMMENT     Notes:
            Streptomyces coelicolor sequencing at The Sanger Centre is funded
            by the BBSRC and Beowulf Genomics
            Details of S. coelicolor sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL; http://www.sanger.ac.uk/Projects/S_coelicolor/)
            CDS are numbered using the following system eg SC7B7.01c. SC (S.
            coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
            strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous.
            The length in codons is given for each CDS.
            Usually the highest scoring match found by fasta -o is given for
            CDS which show significant similarity to other CDS in the database.
            The position of possible ribosome binding site sequences are given
            where these have been used to deduce the initiation codon.
            Gene prediction is based on positional base preference in codons
            using a specially developed Hidden Markov Model (Krogh et al.,
            Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot
            program of Bibb et al., Gene 30:157-66(1984) as implemented at
            http://www.nih.go.jp/
            jun/cgi-bin/frameplot.pl. CAUTION:  We may not have predicted the
            correct initiation codon.  Where possible we choose an initiation
            codon (atg, gtg, ttg or (att)) which is preceded by an upstream
            ribosome binding site sequence (optimally 5-13bp before the
            initiation codon).  If this cannot be identified we choose the most
            upstream initiation codon.
            IMPORTANT: This sequence MAY NOT be the entire insert of the
            sequenced clone.  It may be shorter because we only sequence
            overlapping sections once, or longer, because we arrange for a
            small overlap between neighbouring submissions.
            Cosmid 7H2 Lies between and overlaps with cosmids 3A3 and 3C9 on
            the AseI-C genomic restriction fragment.
FEATURES            Location/Qualifiers
     source         1. .42655
                    /organism="Streptomyces coelicolor A3(2)"
                    /strain="A3(2)"
                    /db_xref="taxon:100226"
                    /clone="cosmid 7H2"
     gene           1. .273
                    /gene="SC7H2.01"
     CDS            <1. .273
                    /gene="SC7H2.01"
                    /note="SC7H2.01, partial CDS, unknown, len: >90aa; similar
                    to the C-terminal region of several hypothetical proteins
                    both prokaryote and eukaryote egs. SW:Y709_SYNY3
                    hypothetical protein from Synechocystis sp. strain PCC6803
                    (130 aa) fasta scores; opt: 100, z-score: 145.9, E():
                    0.84, (34.9% identity in 83 aa overlap) and SW:UK14_CAPHI
                    from Capra hircus (goat) (137 aa) fasta scores; opt: 96,
                    z-score: 140.4, E(): 1.7, (34.6% identity in 81 aa
                    overlap)."
                    /codon_start=1
                    /transl_table=11
                    /label=SC7H2.01
                    /product="hypothetical protein"
                    /protein_id="CAB52043.1"
                    /db_xref="GI:5689880"
                    /translation="DPVRQAEQVLANLTEQLRAVGSGLEHVLSTDVYVVSAETAALSA
                    VWDVVEASGLSTGPHSSTLLGVSCLGYPGQLVEITATAVVPESGRP"
     gene           270. .821
                    /gene="SC7H2.02"
     CDS            270. .821
                    /gene="SC7H2.02"
                    /note="SC7H2.02, possible acetyltransferase, len: 183aa;
                    similar to SW:YIAC_ECOLI hypothetical protein from
                    Escherichia coli (146 aa) fasta scores; opt: 178, z-score:
                    227.9, E(): 2.3e-05, (27.2% identity in 147 aa overlap).
                    Also has region of similarity to several
                    acetyltransferases eg. TR:Q08414 (EMBL:X73149)
                    nourseothricin acetyltransferase from Streptomyces noursei
                    (189 aa) fasta scores; opt: 115, z-score: 150.4, E():
                    0.47, (40.2% identity in 82 aa overlap). Contains Pfam
                    match to entry PF00583 Acetyltransf, Acetyltransferase
                    (GNAT) family."
                    /codon_start=1
                    /transl_table=11
                    /label=SC7H2.02
                    /product="hypothetical protein"
                    /protein_id="CAB52044.1"
                    /db_xref="GI:5689881"
                    /translation="MNREAGARAGADVDAGPSDDAGAGVSTDAGAEAVVLRRARAVDA
                    RAAADVWLRSFAAALPTVVRPRSDSEVRAYFRHVVVERYETWVAQEDDGAVAGVMVLE
                    GEELSQLYLAPERRGCGIGDRFVALAKERRADGLSLWTFQVNAPARRFYERHGFRAVE
                    WTDGDRNEEREPDVRYEWRPRSA"
     misc_feature   375. .740
```

```
                /gene="SC7H2.02"
                /note="Pfam match to entry PF00583 Acetyltransf,
                Acetyltransferase (GNAT) family, score 43.70, E-value
                4.2e-09."
stem_loop       820. .900
                /note="Inverted repeat (Score 57: 23/26 ( 88%) matches, 0
                gaps) with potential 28 base loop."
gene            complement(907. .1878)
                /gene="SC7H2.03c"
CDS             complement(907. .1878)
                /gene="SC7H2.03c"
                /note="SC7H2.03c, probable aminopeptidase, len: 323aa;
                similar to many proline aminopeptidases eg. SW:PIP_XANCI
                proline iminopeptidase from Xanthomonas campestris pv.
                citri (313 aa) fasta scores; opt: 848, z-score: 979.3,
                E(): 0, (44.4% identity in 320 aa overlap). Contains Pfam
                match to entry PF00561 abhydrolase, alpha/beta hydrolase
                fold."
                /codon_start=1
                /transl_table=11
                /label=SC7H2.03c
                /product="putative aminopeptidase"
                /protein_id="CAB52045.1"
                /db_xref="GI:5689882"
                /translation="MSLYPEIEPYDHGMLDVGDGNHVYWETCGNPHGKPAVVLHGGPG
                SRASPGLRRYFDPAAYRIVLLDQRGAGRSLPRASAPDTDMSVNTTAHLMADLERLRVH
                LGIERWLVWGVSWGSVLGLRYAQTHPGVVTELVLTGVATGSNAEVALLTRGLGNIFPE
                AHERFLAELPPDARDGNLPAAYNRLLESPDPAVRERAARAWTDWETATIPAPPGSVAR
                YQDPDFRMGFARTVTHYWGNDHFLGDGNDEGVVIRDAHLLKGIPGTLVQGSLDFGNLL
                GIVWRLHHAWPDSDLVIVDEAGHDAGTTGDEALLAATDKYARGGTAE"
misc_feature    complement(928. .1701)
                /gene="SC7H2.03c"
                /note="Pfam match to entry PF00561 abhydrolase, alpha/beta
                hydrolase fold, score 74.20, E-value 2.7e-18."
gene            complement(1933. .2646)
                /gene="SC7H2.04c"
CDS             complement(1933. .2646)
                /gene="SC7H2.04c"
                /note="SC7H2.04c, unknown, len: 237aa; similar to
                SW:Y08N_MYCTU hypothetical protein from Mycobacterium
                tuberculosis (350 aa) fasta scores; opt: 870, z-score:
                990.4, E(): 0, (57.1% identity in 231 aa overlap)."
                /codon_start=1
                /transl_table=11
                /label=SC7H2.04c
                /product="hypothetical protein"
                /protein_id="CAB52046.1"
                /db_xref="GI:5689883"
                /translation="MDTGGLSVLDRRIEGCRACPRLVEWREEVARTKRAAFADWTYWG
                RPVPGFGPPDARLLIVGLAPAAHGGNRTGRMFTGDRSGDVLYQALYDVGLASQPTAVR
                VDDGLELYGVRVTSPVHCAPPANKPTPAERDTCRSWLVQELGLLRPTLRAVVVLGAFG
                WQAALPAFAGAGWTVPRPRPAFAHGTQVTLDAADGPDLHLFGCFHVSQRNTFTGRLTP
                EMLRDVLRTAAETAGLPAR"
gene            2763. .3323
                /gene="SC7H2.05"
CDS             2763. .3323
                /gene="SC7H2.05"
                /note="SC7H2.05, unknown, len: 186aa; similar to
                SW:YRFH_ECOLI hypothetical protein from Escherichia coli
                (133 aa) fasta scores; opt: 228, z-score: 249.5, E():
                1.4e-06, (35.2% identity in 122 aa overlap)."
                /codon_start=1
                /transl_table=11
                /label=SC7H2.05
                /product="hypothetical protein"
                /protein_id="CAB52047.1"
                /db_xref="GI:5689884"
                /translation="MASEHDEGGSGPVAAPRDGAAAAAGTPDPKTATPDPKTAAAIAA
                AEAAGPQHGESVRVDSWIWAVRLIKTRSLGATACRGGHVRVNGERVKPAYSLRVGDEV
                RLRQEGRERVVVVKRLIRKRVGAPVAVQCYVDNSPPPPPREAVAPAGIRDRGAGRPTK
                RDRREMDRLRGLEGLSSSRRDAGTRP"
gene            complement(3375. .3881)
                /gene="SC7H2.06c"
CDS             complement(3375. .3881)
                /gene="SC7H2.06c"
                /note="SC7H2.06c, possible integral membrane protein, len:
                168aa; contains possible membrane spanning hydrophobic
                regions."
                /codon_start=1
                /transl_table=11
                /label=SC7H2.06c
                /product="putative integral membrane protein"
                /protein_id="CAB52048.1"

  Query Match             4.4%;  Score 72.8;  DB 64;  Length 42655;
  Best Local Similarity   46.7%;  Pred. No. 0.019;
  Matches  230;  Conservative    0;  Mismatches  262;  Indels    0;  Gaps    0;

Qy      66 gagcaggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgcttccgg 125
           |||| ||||  | || |||||  | | |    ||   |  |||||  | |||||  | |
Db   37276 GAGCTGGCGAACGACCCGAGCGCCACCGGCCAGGAGGGTGCCGCCCGCTGCCGCGCCGCG 37217

Qy     126 ccgcccaccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcg 185
           | |||   |  | |  ||  |      ||||||    ||   || ||| ||||
Db   37216 CTGCCGGACGCCGAGGCCAAGGCCCGCGCCTGGGAGGCCATGTTCGCCTCCGACGACCTC 37157

Qy     186 aggtaccccaggtacctgcctgaagcctactgcctgtgccgggggctgcctgaccgggctg 245
              ||  |  || |    |    ||| |  || | || | |  |   | | ||| ||||
Db   37156 TCCAACTACCTGTTCACCGCCACCGCCCAGGGCTTCTGGCAGCCCGAACAGGCCGAGCTG 37097

Qy     246 ttcggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctg 305
           || | ||  | |   ||||| |  |  ||   | |||    |  || ||  |  |
Db   37096 GTCCGGGACTACGTACCGCGCTACTACCCCGAGGCCGTCGCGCTCGCCGCCCGCCGCGGT 37037

Qy     306 cgccgcaccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatc 365
           | |  ||   ||| |  |||||| || |  || ||  | || | ||   || | ||  |
Db   37036 CCCGCCATGGCCGACGCCGCCGGCCGCTGGGCCTTCCCCGCCCACGCGGTCGACGCCGAC 36977

Qy     366 cccgtgggctgcacctgcgtccccgagccggagaaggacgcagacagcatcaactccagc 425
           || || ||     | | |    |   | ||   ||  |   |  |  |     | ||
Db   36976 ACCCTGCGCCTGGGCCGGGAGTGCCTCGCCGACGCCGATCCGATCCCCGCCCTGCGCCGC 36917

Qy     426 atcgacaaacagggcgccaagctcctgctgggccccaacgacgcgcccgctggcccctga 485
           |    ||   |      | |   |||| | || ||      |  |  |  |  | |
Db   36916 AAGCTCACCGACCAGCTCGACGACCTGGCGCGCGCCCTGCGGGTACGGGAGGCGCACACG 36857

Qy     486 ggccggtcctgccccgggaggtctccccggcccgcatcccgaggcgcccaagctggagcc 545
           | | |  || ||| |||  || | || || || ||    |  | ||  |   | | |
Db   36856 GACTGACCCGGCCGCGGCCGGCCCCCGCGCCCGGCGGGGCCGGCCGTGCCCCCGGTGGGG 36797

Qy     546 gcctggagggct 557
             | |  ||| |
Db   36796 AGCCGCCGGGAT 36785


RESULT  12
D76415
LOCUS       D76415       1429 bp    mRNA            PLN       01-FEB-2000
DEFINITION  Rice seedling mRNA for cysteine proteinase, complete cds.
ACCESSION   D76415
VERSION     D76415.1  GI:1514952
KEYWORDS    cysteine proteinase.
SOURCE      Oryza sativa (strain:cv. Koshihikari) seedling aleurone layer and
            endosperm cDNA to mRNA, clone:pREP1.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
            Poaceae; Oryza.
REFERENCE   1  (bases 1 to 1429)
  AUTHORS   Kato,H. and Minamikawa,T.
  TITLE     Identification and characterization of a rice cysteine
            endopeptidase that digests glutelin
  JOURNAL   Eur. J. Biochem. 239 (2), 310-316 (1996)
  MEDLINE   96314485
REFERENCE   2  (bases 1 to 1429)
  AUTHORS   Katoh,H.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-1995) to the DDBJ/EMBL/GenBank databases. Hideki
            Katoh, Tokyo Metropolitan University, Department of Biology;
            Minami-osawa 1-1, Hachioji, Tokyo 192-03, Japan (Tel:0426-77-2562,
            Fax:0426-77-2559)
FEATURES             Location/Qualifiers
     source          1. .1429
                     /organism="Oryza sativa"
                     /strain="cv. Koshihikari"
                     /db_xref="taxon:4530"
                     /clone="pREP1"
                     /dev_stage="seedling"
                     /tissue_type="aleurone layer and endosperm"
     5'UTR           1. .19
     sig_peptide     20. .103
                     /product="putative"
     CDS             20. .1126
                     /codon_start=1
                     /evidence=experimental
                     /product="cysteine proteinase"
                     /protein_id="BAA11170.1"
                     /db_xref="GI:1514953"
                     /translation="MGRVISSWRVLAVVAALMAMAAVELCAAIPFDERDLESDEALWD
                     LYERWQEHHHVPRHHGEKHRRFGAFKDNVRYIHEHNKRAPGYAPLNRFGDMGREEFRA
                     TFAGSHANDLRRDGLAAPPLPGFMYEGVRDLPRAVDWRRKGAVTGVKDQGKCGSCWAF
                     STVVSVEGINAIRTGRLVSLSEQELIDCDTADNSGCQGGLMENAFEYIKHSGGITTES
                     AYPYRAANGTCDAVRARGGLVVIDGHQNVPANSEAALAKAVANQPVSVAIDAGDQSFQ
                     FYSDGVFAGDCGTDLDHGVAVVGYGETNDGTEYWIVKNSWGTAWGEGGYIRMQRDSGY
                     DGGLCGIAMEASYPVKFSPNRVTPRRALGPKETQ"
     mat_peptide     413. .1123
                     /product="cysteine proteinase"
     3'UTR           1127. .1429
     polyA_signal    1401. .1410
                     /note="putative"
     polyA_site      1429
BASE COUNT      259 a    441 c    471 g    258 t
ORIGIN


  Query Match             4.0%;  Score 65;  DB 7;  Length 1429;
  Best Local Similarity   45.2%;  Pred. No. 0.59;
  Matches  239;  Conservative    0;  Mismatches  290;  Indels    0;  Gaps    0;

Qy      28 tcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaacgcgagct 87
           |||  ||  |||       || || | ||  || |   ||||||| ||  | ||   ||
Db     525 TCAACGCGATCCGGACGGGGCGGCTGGTGTCGCTGTCGGAGCAGGAGCTGATCGACTGCG 584

Qy      88 gcccggcaggggcaggcccgccgaccgccgcttccggccgcccaccaacctgcgcagcg 147
             | |||| |    |||   | | | || |  ||  |  | | |   |  | || |
Db     585 ACACGGCGGACAACAGCGGCTGCCAGGGCGGGCTCATGGAGAACGCGTTCGAGTACATCA 644

Qy     148 tgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtacctgcctg 207
             |     | | | |  ||  |    | || |    ||     | ||    ||    |
Db     645 AGCACAGCGGCGGCATCACCACCGAGTCCGCCTACCCGTACCGCGCCGCCAACGGAACGT 704

Qy     208 aagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggacgtgcgct 267
             | |  |  ||     || ||| | |||   | | |   |||  |  || ||||||
Db     705 GCGACGCCGTCCGCGCGCGCGGCGGGCTGGTGGTGATCGACGGGCACCAGAACGTGCCGG 764

Qy     268 tccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcctgcgccg 327
             |  |||||   | |    |       ||||| |   |  | |  | ||| |  |  ||
Db     765 CCAACAGCGAGGCCGCGCTCGCCAAGGCCGTCGCCAACCAGCCCGTCTCCGTCGCCATCG 824

Qy     328 gcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacctgcgtcc 387
             || | |   | | |  || |||  ||||  |  |  | | | | ||  ||||| |
Db     825 ACGCCGGCGACCAGTCCTTCCAGTTCTACTCCGACGGCGTCTTCGCCGGCGACTGCGGCA 884

Qy     388 ccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgccaagc 447
             |||| |  ||  | | ||  |      ||  || | ||     |||  | ||| |  ||
Db     885 CCGACCTCGACCACGGCGTCGCGGTGGTCGGCTACGGCGAGACCAACGACGGCACGGAGT 944

Qy     448 tcctgctgggccccaacgacgcgcccgctggcccctgaggccggtcctgccccgggaggt 507
             | | | | |   |||     | | | | |    | |   |   || |  | ||| |
```

```
Db     945 ACTGGATCGTCAAGAACTCGTGGGGCACAGCCTGGGGCGAGGGCGGCTACATCCGGATGC 1004

Qy     508 ctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggc 556
             | |    |||  | |   |||| ||     ||   ||||  |   ||
Db    1005 AGCGCGACTCCGGCTACGACGGCGGCCTCTGCGGCATCGCCATGGAAGC 1053


RESULT  13
AF099203
LOCUS       AF099203     2588 bp    DNA             PLN       07-FEB-2000
DEFINITION  Oryza sativa cysteine endopeptidase precursor (EP3A) gene, complete
            cds.
ACCESSION   AF099203
VERSION     AF099203.1  GI:4426616
KEYWORDS    .
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
REFERENCE   1  (bases 1 to 2588)
  AUTHORS   Ho,S.L., Tong,W.F. and Yu,S.M.
  TITLE     Multiple mode regulation of a cysteine proteinase gene expression
            in rice
  JOURNAL   Plant Physiol. 122 (1), 57-66 (2000)
  MEDLINE   20098713
   PUBMED   10631249
REFERENCE   2  (bases 1 to 2588)
  AUTHORS   Ho,S.L. and Tong,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-1998) Department of Biology, National Taiwan
            Normal University, Taipei, Taiwan 11718, Republic of China
FEATURES             Location/Qualifiers
     source          1. .2588
                     /organism="Oryza sativa"
                     /db_xref="taxon:4530"
     gene            818. .>2080
                     /gene="EP3A"
     TATA_signal     818. .825
                     /gene="EP3A"
     mRNA            852. .>2080
                     /gene="EP3A"
                     /product="cysteine endopeptidase precursor"
     sig_peptide     974. .1057
                     /gene="EP3A"
     CDS             974. .2080
                     /gene="EP3A"
                     /codon_start=1
                     /product="cysteine endopeptidase precursor"
                     /protein_id="AAD20453.1"
                     /db_xref="GI:4426617"
                     /translation="MGRVISSWRVLAVVAALMAMAAVELCAAIPFDERDLESDEALWD
                     LYERWQEHHHVPRHHGEKHRRFGAFKDNVRYIHEHNKRAPGYPPLNRFGDMGREEFRA
                     TFAGSHANDLRRDGLAAPPLPGFMYEGVRDLPRAVDWRRKGAVTGVKDQGKCGSCWAF
                     STVVSVEGINAIRTGRLVSLSEQELIDCDTADNSGCQGGLMENAFEYIKHSGGITTES
                     AYPYRAANGTCDAVRARGGLVVIDGHQNVPANSEAALAKAVANQPVSVAIDAGDQSFQ
                     FYSDGVFAGDCGTDLDHGVAVVGYGETNDGTEYWIVKNSWGTAWGEGGYIRMQRDSGY
                     DGGLCGIAMEASYPVKFSPNRVTPRRALGPKETQ"
     mat_peptide     1058. .2077
                     /gene="EP3A"
                     /product="cysteine endopeptidase"
     misc_feature    1649. .1657
                     /gene="EP3A"
                     /note="glycosylation site"
BASE COUNT      607 a    745 c    682 g    554 t
ORIGIN


  Query Match             4.0%;  Score 65;  DB 6;  Length 2588;
  Best Local Similarity   45.2%;  Pred. No. 0.53;
  Matches  239;  Conservative    0;  Mismatches  290;  Indels    0;  Gaps    0;

Qy      28 tcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaacgcgagct 87
           |||  ||  |||       || || | ||  || |   ||||||| ||  | ||   ||
```

```
Db   1479 TCAACGCGATCCGGACGGGGCGGCTGGTGTCGCTGTCGGAGCAGGAGCTGATCGACTGCG 1538

Qy     88 gcccggcaggggcaggcccgccgaccgccgcttccggccgcccaccaacctgcgcagcg 147

Db   1539 ACACGGCGGACAACAGCGGCTGCCAGGGCGGGCTCATGGAGAACGCGTTCGAGTACATCA 1598

Qy    148 tgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtacctgcctg 207

Db   1599 AGCACAGCGGCGGCATCACCACCGAGTCCGCCTACCCGTACCGCGCCGCCAACGGAACGT 1658

Qy    208 aagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggacgtgcgct 267

Db   1659 GCGACGCCGTCCGCGCGCGCGGCGGGCTGGTGGTGATCGACGGGCACCAGAACGTGCCGG 1718

Qy    268 tccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcaccccgcctgcgccg 327

Db   1719 CCAACAGCGAGGCCGCGCTCGCCAAGGCCGTCGCCAACCAGCCCGTCTCCGTCGCCATCG 1778

Qy    328 gcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacctgcgtcc 387

Db   1779 ACGCCGGCGACCAGTCCTTCCAGTTCTACTCCGACGGCGTCTTCGCCGGCGACTGCGGCA 1838

Qy    388 ccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgccaagc 447

Db   1839 CCGACCTCGACCACGGCGTCGCGGTGGTCGGCTACGGCGAGACCAACGACGGCACGGAGT 1898

Qy    448 tcctgctgggccccaacgacgcgcccgctggcccctgaggccggtcctgccccgggaggt 507

Db   1899 ACTGGATCGTCAAGAACTCGTGGGGCACAGCCTGGGGCGAGGGCGGCTACATCCGGATGC 1958

Qy    508 ctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggc 556

Db   1959 AGCGCGACTCCGGCTACGACGGCGGCCTCTGCGGCATCGCCATGGAAGC 2007
```

```
RESULT  14
AF005277
LOCUS       AF005277     3578 bp    DNA             BCT       02-MAR-1998
DEFINITION  Cellulomonas biazotea cellobiase (cba) gene, complete cds.
ACCESSION   AF005277
VERSION     AF005277.1  GI:2921739
KEYWORDS    .
SOURCE      Cellulomonas biazotea.
  ORGANISM  Cellulomonas biazotea
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Micrococcineae; Cellulomonadaceae; Cellulomonas.
REFERENCE   1  (bases 1 to 3578)
  AUTHORS   Wong,W.K., Ali,A., Chan,W.K., Ho,V. and Lee,N.T.
  TITLE     The cloning, expression and characterization of a cellobiase gene
            encoding a secretory enzyme from Cellulomonas biazotea
  JOURNAL   Gene 207 (1), 79-86 (1998)
  MEDLINE   98172737
REFERENCE   2  (bases 1 to 3578)
  AUTHORS   Wong,W.-K.R., Ali,A., Chan,R.W.K., Ho,V. and Lee,N.T.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-1997) Biochemistry, The Hong Kong University of
            Science and Technology, Clear Water Bay Road, Kowloon, Hong Kong
FEATURES             Location/Qualifiers
     source          1. .3578
                     /organism="Cellulomonas biazotea"
                     /strain="ATCC 486"
                     /db_xref="taxon:1709"
                     /db_xref="ATCC:486"
     gene            327. .2813
                     /gene="cba"
     CDS             327. .2813
                     /gene="cba"
                     /function="hydrolyzing beta-D-glucoside and cellobiose"
                     /note="one type of cellulase"
                     /codon_start=1
                     /transl_table=11
                     /product="cellobiase"
                     /protein_id="AAC38196.1"
                     /db_xref="GI:2921740"
                     /translation="MTSQTALDPAALVASLPLETKVRLLTGATAFTLAPEESIGLGEV
                     RLSDGPTGVRGLKFSGGRTVALFPNATLLASAWSEESTTEVGRLLAEEALAQQIHVVL
                     GPTINLHRSVLGGRLFEAYSEDPLLTGRLAAAYVRGLQDLGVGACLKHLVANESETER
                     NTMNSVVDPATLRELYLLPFEIAVDESDPWSVMAAYNDVNGVPATEHHHVVNEVLKGE
                     WGYTGLVMSDWFATRTAAPAAAGGLDLVMPGPDGPWGDALVAAVRSGELDESVVDDHL
                     RRLLVLAARVGALGDLRDYPDDLPAPDSAVRREQLTRLAAAGMTVLTNADDTLPLARG
                     TRVALVGRHALETIDMGGGSATVNPPYQVSVAEGLTALLGDAVDVVDGVEVRTRPVPA
                     RPGFVVDPDTGRPGLHLTLLAADGTVLDERHDAPSTVMVGFDDDFPQAVARVRFRARV
                     AGEGALEVGAIGVGRWQVTAGGTELAWTLATSGTGFAEEMLAPPTRTDQVHVGSDAVV
                     DATVVLRSSTRSVTVGDADPGTDAGAAAEPLAGVGLFGLVARPAPEAEDDVITRAAAA
                     AAQADVAVVVVGLTEEEETESVDKSTIALPGAQDALVRAVAAAARRTVVVVNAATPVL
                     MPWLDDVDAVLWAGLPGQEGGHAVAAALLGDQEPTGRLVTTFPAADGAAPAWSVTPVD
                     GDLEYTEGRFVGYRGHWADRAPAPAFWLGHGLGYATWEYADATLDTDGDAPAVTVTVT
                     NTGARTSREVVQVYLEPASSDEPVRLVGWADATVDAGASARVTVTADARMWRRWDEAA
                     GGWSRLADGGRLLVARGLGDVRATLALPTA"
BASE COUNT      426 a   1421 c   1306 g    425 t
ORIGIN
```

```
  Query Match             3.9%;  Score 63.6;  DB 1;  Length 3578;
  Best Local Similarity   48.4%;  Pred. No. 0.81;
  Matches  177;  Conservative    0;  Mismatches  189;  Indels    0;  Gaps    0;

Qy    189 taccccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttc 248

Db    687 TACTCCGAGGACCCGCTGCTCACCGGCCGGCTCGCCGCCGCCTACGTGCGCGGCCTGCAG 746

Qy    249 ggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgc 308

Db    747 GACCTCGGCGTCGGCGCCTGCCTCAAGCACCTCGTCGCCAACGAGTCGGAGACCGAGCGC 806

Qy    309 cgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcccc 368

Db    807 AACACCATGAACTCCGTCGTCGACCCCGCGACGCTCCGCGAGCTCTACCTGCTGCCGTTC 866

Qy    369 gtgggctgcacctgcgtccccgagccggagaaggacgcagacagcatcaactccagcatc 428

Db    867 GAGATCGCCGTCGACGAGTCCGACCCGTGGTCCGTCATGGCCGCGTACAACGACGTCAAC 926

Qy    429 gacaaacagggcgccaagctcctgctgggccccaacgacgcgcccgctggcccctgaggc 488

Db    927 GGCGTCCCCGCGACCGAGCACCACCACGTCGTCAACGAGGTCCTCAAGGGCGAGTGGGGC 986

Qy    489 cggtcctgccccgggaggtctccccggcccgcatcccgaggcgcccaagctggagccgcc 548

Db    987 TACACCGGCCTCGTCATGTCCGACTGGTTCGCGACCCGGACCGCCGCACCCGCCGCCGCC 1046

Qy    549 tggagg 554

Db   1047 GGCGGG 1052
```

```
RESULT  15
AC025783
LOCUS       AC025783   180630 bp    DNA             HTG       27-JUL-2000
DEFINITION  Oryza sativa chromosome 10 clone OSJNBa0001O14, *** SEQUENCING IN
            PROGRESS ***, 1 ordered pieces.
ACCESSION   AC025783
VERSION     AC025783.5  GI:9502436
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
REFERENCE   1  (bases 1 to 180630)
  AUTHORS   Buell,R., Benito,M.-I., Hsiao,J., Zismann,V., Sellers,P., Pai,G.,
            Militscher,J., Vanaken,S., Bowman,C., Craven,B., Moffat,K.,
            Fujii,C., Hill,J., Jenkins,C., Khalak,H., Feldblyum,T.,
            Quackenbush,J., White,O., Salzberg,S. and Fraser,C.
  TITLE     Oryza sativa ssp. japonica cv. Nipponbare OSJNBa0001O14 BAC genomic
            sequence
  JOURNAL   Unpublished
```

```
REFERENCE   2  (bases 1 to 180630)
  AUTHORS   Buell,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     On Jul 27, 2000 this sequence version replaced gi:9440586.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   180630: contig of 180630 bp in length.
FEATURES            Location/Qualifiers
     source         1. .180630
                    /organism="Oryza sativa"
                    /db_xref="taxon:4530"
                    /chromosome="10"
                    /clone="OSJNBa0001O14"
BASE COUNT    51234 a  38444 c  37306 g  53639 t      7 others
ORIGIN


  Query Match             3.9%;  Score 63.6;  DB 49;  Length 180630;
  Best Local Similarity   46.8%;  Pred. No. 0.37;
  Matches  232;  Conservative    1;  Mismatches  260;  Indels    3;  Gaps    1;

Qy      41 ccacacgctgcagctggggccgcgtgagcaggcgcgcaacgcgagctgcccggcaggggg 100
           ||| |   ||     || |||||| | |  |   |    |||   |  |  |     ||
Db 131327 CCAGAGCATGGCCACGGCGCCGCGCGCGAGGCGCCTGGCCGCCGCCGACGTGCAGCCGGC 131386

Qy     101 caggcccgccgaccgccgcttccggccgcccaccaacctgcgcagcgtgtcgccctgggc 160
           || ||| || | | |||                    || |  ||:| | ||| |  | ||
Db 131387 CACGCCGGCGGTCGGCCCAGCGTCCGGCGGGCGGCCCCAGTTCARCCTCTCGTCGGGCGC 131446

Qy     161 ctacagaatctcctacgacccggcgaggtaccccaggtacctgcctgaagcctactgcct 220
              ||    |       |    | |    || | | | | | |       |  || |
Db 131447 GGCCACCGCCGTGGTATTCGTCTCCATCGTCCTCTGCTTCATCCTGCTCTGCACCTACTG 131506

Qy     221 gtgccggggctgcctgaccgggctgttcggcgaggaggacgtgcgcttccgcagcgcccc 280
             || | | | | | |  ||   |   ||||| |     |     | | |||     ||
Db 131507 CCGCTGCGCCAGGCAGCGCGCCATCGCCGGCGCGCGCAGCCGCGTCATGCGCGAGATCCG 131566

Qy     281 tgtctacatgcccaccgtcgtcctgcgccgcacccccgcctgcgccggcggccgttccgt 340
            |    | | ||   |||| |||| | |||| || |||||  |||||  ||  |   ||||
Db 131567 GGAGCGCGTCCCGGGCGTCCTCCTCCTCCGCCCCGCCGCCGCCGCCGCCGCGCTGCCCGT 131626

Qy     341 ctacaccgaggcctacgtcaccatccccgtgggctgcacctgcgtccccgagccggagaa 400
           |  | || |  ||  || | ||  | ||| ||||  ||      |  ||      ||| |
Db 131627 CCTCCCCTACTCCGCCGCCGCCGCCGCCGCGGGCGCCAAGAAGGGGCCGCTCGTCGAGGA 131686

Qy     401 ggacgcagacagcatcaactccagcatcgacaaacagggcgccaagctcctgc---tggg 457
              | | | || || |  ||  |  || | |  | |||| |||| || | |    | |
Db 131687 CTGCCCCGTCTGCCTCGAGGCCTTCGCCGGCGACGACGGCGTCAAGGTCGTCCCCGCGTG 131746

Qy     458 ccccaacgacgcgcccgctggcccctgaggccggtcctgccccgggaggtctccccggcc 517
           |  | |||     | ||| | |  |   | || ||    | ||| | ||    || | ||
Db 131747 CGGCCACGTGTTCCACGCCGCCTGCATCGACCAGTGGCTCGCCGTGCGGAACTCCTGCCC 131806

Qy     518 cgcatcccgaggcgcc 533
           ||  | |||  |||||
Db 131807 CGTCTGCCGCTGCGCC 131822


Search completed: February 24, 2001, 20:48:02
Job time: 8990 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 21:04:27 ; Search time 1413.06 Seconds
(without alignments)
8142.821 Million cell updates/sec

Title: US-09-320-713-3
Perfect score: 1642
Sequence: 1 ggaattcggcacgagctcgt.....tgaaaaaaaaaaaaaaaaaaa 1642

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*
44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estpl1:*
70: em_estpl2:*
71: em_estpl3:*
72: em_estpl4:*
73: em_estpl5:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estpl6:*
118: em_estpl7:*
119: em_estpl8:*
120: em_estro14:*
121: em_estro15:*
122: em_estro16:*
123: em_estro17:*
124: em_estro18:*
125: em_estro19:*
126: gb_est58:*
127: gb_est59:*
128: gb_est60:*
129: gb_est61:*
130: gb_est62:*
131: gb_est63:*
132: gb_est64:*
133: gb_est65:*
134: gb_est66:*
135: gb_est75:*
136: gb_est76:*
137: gb_est77:*
138: gb_est78:*
139: gb_est79:*
140: gb_est80:*
141: gb_est81:*
142: gb_est82:*
143: gb_est83:*
144: gb_est84:*
145: gb_est85:*
146: gb_est86:*
147: gb_est87:*
148: gb_gss1:*
149: gb_gss2:*
150: gb_gss3:*
151: gb_gss4:*
152: em_gss1:*
153: em_gss2:*
154: em_gss3:*
155: em_gss4:*
156: gb_gss5:*
157: gb_gss6:*
158: gb_gss7:*
159: gb_gss8:*
160: gb_gss9:*
161: em_gss5:*
162: em_gss6:*
163: em_gss7:*
164: em_gss8:*
165: em_gss9:*
166: em_gss10:*
167: em_gss11:*
168: gb_gss10:*
169: gb_gss11:*
170: em_gss12:*
171: gb_gss12:*
172: gb_gss13:*
173: gb_gss14:*
174: gb_gss15:*
175: gb_gss16:*
176: gb_gss17:*
177: gb_gss18:*
178: gb_gss19:*
179: em_gss13:*
180: gb_gss20:*
181: gb_gss21:*
182: gb_gss22:*
183: gb_gss23:*
184: gb_gss24:*
185: em_gss14:*
186: em_gss15:*
187: em_gss16:*
188: em_gss17:*
189: em_gss18:*
190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	579.6	35.3	586	136	BE856748	BE856748 7f68b12.x
2	545.4	33.2	579	97	AW956791	AW956791 EST368861
c 3	527	32.1	529	26	AI870335	AI870335 wl72f12.x
c 4	517.2	31.5	522	108	BE465289	BE465289 hw14c08.x
c 5	480.6	29.3	503	38	AW015888	AW015888 UI-H-BI0-
c 6	467	28.4	476	25	AI817213	AI817213 wg62h12.x
c 7	456.2	27.8	461	11	AA778133	AA778133 zf46h06.s
c 8	452	27.5	487	21	AI492261	AI492261 tf59b03.x
c 9	448.8	27.3	466	143	N52801	N52801 yz18g05.s1
c 10	448.2	27.3	453	93	AW614686	AW614686 hh31b08.x
c 11	442.4	26.9	446	22	AI569688	AI569688 tn04e05.x
c 12	439.8	26.8	443	17	AI199159	AI199159 qi46h07.x
c 13	421.8	25.7	475	143	N92873	N92873 zb71d08.s1
c 14	420.2	25.6	425	19	AI371173	AI371173 ta09c11.x
c 15	417	25.4	419	16	AI085287	AI085287 qf18f01.x
c 16	413.8	25.2	425	16	AI146504	AI146504 qb84h09.x
c 17	411.8	25.1	469	141	H29012	H29012 ym31h04.s1
c 18	410.2	25.0	425	24	AI740658	AI740658 wg07a09.x
c 19	408.6	24.9	427	87	AW236638	AW236638 xm48h05.x
c 20	407.6	24.8	426	17	AI206346	AI206346 qr21h05.x
21	405.6	24.7	414	141	F36829	F36829 HSPD34729 H
22	398.4	24.3	441	18	AI275406	AI275406 ql63c10.x
c 23	396.4	24.1	406	19	AI339872	AI339872 qq43d12.x
c 24	395	24.1	465	141	H30762	H30762 yo79b08.r1
c 25	393	23.9	408	17	AI168396	AI168396 qa25d04.s
c 26	391.8	23.9	407	23	AI634093	AI634093 wa87g06.x
c 27	391	23.8	395	16	AI093060	AI093060 qa96f08.x
c 28	385.8	23.5	392	8	AA478614	AA478614 zv19c10.s
c 29	384	23.4	398	16	AI095032	AI095032 qa18g04.x
c 30	363	22.1	484	141	H14168	H14168 ym62g01.r1
31	348.4	21.2	412	5	AA303602	AA303602 EST16312
32	325.2	19.8	434	146	W25618	W25618 zb71d08.r1
c 33	324.2	19.7	361	16	AI144123	AI144123 qb62b08.x
c 34	320	19.5	341	18	AI277547	AI277547 ql97c06.x
c 35	300	18.3	375	142	N34182	N34182 yx77c03.r1
c 36	277.8	16.9	307	15	AI074646	AI074646 ox81g11.s
c 37	275.6	16.8	481	142	N20280	N20280 yx42f05.s1
c 38	274	16.7	323	23	AI675296	AI675296 wc21f06.x
39	256.8	15.6	263	140	F17157	F17157 HSPD01678 H
40	255.6	15.6	269	6	AA385051	AA385051 EST98645
c 41	233	14.2	235	19	AI383176	AI383176 tc44f05.x
42	226.2	13.8	239	141	F35345	F35345 HSPD31416 H
43	225.2	13.7	254	141	F32964	F32964 HSPD26092 H
44	165.2	10.1	183	143	N57623	N57623 yz18g05.r1
45	145	8.8	535	19	AI357592	AI357592 qy14f05.x

ALIGNMENTS

RESULT 1
BE856748/c
LOCUS BE856748 586 bp mRNA EST 29-SEP-2000
DEFINITION 7f68b12.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:3299807 3', mRNA sequence.
ACCESSION BE856748
VERSION BE856748.1 GI:10370087
KEYWORDS EST.
SOURCE human.

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 586)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .586
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3299807"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT      163 a    140 c    121 g    162 t
ORIGIN
```

```
  Query Match             35.3%;  Score 579.6;  DB 136;  Length 586;
  Best Local Similarity   99.3%;  Pred. No. 2e-132;
  Matches  582;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   1056 gataaattttgtagctggtacactctggcctgggtctctgaattcagcctgtcaccgatg 1115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 GATAAATTTTGTAGCTGGTACACTCTGGCCTGGGTCTCTGAATTCAGCCTGTCACCGATG 527

Qy   1116 gctgactgatgaaatggacacgtctcatctgacccactcttccttccactgaaggtcttc 1175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 GCTGACTGATGAAATGGACACGTCTCATCTGACCCACTCTTCCTTCCACTGAAGGTCTTC 467

Qy   1176 acgggcctccaggtggaccaaagggatgcacaggcggctcgcatgccccagggccagcta 1235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 ACGGGCCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTA 407

Qy   1236 agagttccaaagatctcagatttggttttagtcatgaatacataaacagtctcaaactcg 1295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 AGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCG 347

Qy   1296 cacaatttttttccccctttgaaagccactggggccaatttgtggttaagaggtggtgag 1355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 CACAATTTTTTCCCCCTTTGAAAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGTGAG 287

Qy   1356 ataagaagtggaacgtgacatctttgccagttgtcagaagaatccaagcaggtattggct 1415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 ATAAGAAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCT 227

Qy   1416 tagttgtaagggctttaggatcaggctgaatatgaggacaaagtgggccacgttagcatc 1475
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    226 TAGTTGTAAGGGCTTTAGGATCAGGCCGAATATGAGGACAAAGTGGGCCACGTTAGCATC 167
```

```
Qy   1476 tgcagagatcaatctggaggcttctgtttctgcattctgccacgagagctaggtccttga 1535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 TGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGA 107

Qy   1536 tcttttctttagattgaaagtctgtctctgaacacaattatttgtaaaagttagtagttc 1595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    106 TCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATTTGTAAAAGTTAGAAGTTC 47

Qy   1596 ttttttaaatcattaaaagaggcttgctgaaaaaaaaaaaaaaaaa 1641
          |||||||||||||||||||||||||||||  ||||||||||||
Db     46 TTTTTTAAATCATTAAAAGAGGCTTGCTGAAGGAAAAAAAAAAAA 1
```

```
RESULT   2
AW956791
LOCUS       AW956791      579 bp    mRNA            EST       01-JUN-2000
DEFINITION  EST368861 MAGE resequences, MAGD Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW956791
VERSION     AW956791.1  GI:8146474
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 579)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 96
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .579
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGD"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      168 a    117 c    121 g    171 t      2 others
ORIGIN
```

```
  Query Match             33.2%;  Score 545.4;  DB 97;  Length 579;
  Best Local Similarity   98.4%;  Pred. No. 5.5e-124;
  Matches  560;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy    872 aagcaatctaaaaataataataagtatagcgactatatacctactttaaaatcaactgt 931
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAGCAATCTAAAAATAATAATAAGTATAGCGACTATATACCTACTTTTAAAATCAACTGT 60

Qy    932 tttgaatagaggcagagctattttatattatcaaatgagagctactctgttacatttctt 991
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTGAATAGAGGCAGAGCTATTTTATATTATCAAATGAGAGCTACTCTGTTACATTTCTT 120

Qy    992 aacatataaacatcgttttttacttcttctggtagaatttttttaaagcataattggaatc 1051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACATATAAACATCGTTTTTTACTTCTTCTGGTAGAATTTTTTAAAGCATAATTGGAATC 180

Qy   1052 cttggataaattttgtagctggtacactctggcctgggtctctgaattcagcctgtcacc 1111
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTTGGATAAATTTTGTAGCTGGTACACTCTGGCCTGGGTCTCTGAATTCAGCCTGTCACC 240

Qy   1112 gatggctgactgatgaaatggacacgtctcatctgacccactcttccttccactgaaggt 1171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GATGGCTGACTGATGAAATGGACACGTCTCATCTGACCCACTCTTCCTTCCACTGAAGGT 300
```

```
Qy   1172 cttcacgggcctccaggtggaccaaagggatgcacaggcggctcgcatgccccagggcca 1231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTTCACGGGCCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCA 360

Qy   1232 gctaagagttccaaagatctcagatttggttttagtcatgaatacataaacagtctcaaa 1291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCTAAGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAA 420

Qy   1292 ctcgcacaatttttttccccctttttgaaagccactggggccaatttgtggttaagaggtgg 1351
          |||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db    421 CTCGCACAATTTTTTTCCCCCTTTTGAAAGCCACTGGGGCCAAATTGTGGTT-AGAGGTGG 479

Qy   1352 tgagataagaagtggaacgtgacatctttgccagttgtcagaagaatccaagcaggtatt 1411
          |||||||||||||||||||||||||||||||   |||||||||||||||| ||||||||
Db    480 TGAGATAAGAAGTGGAACGTGACATCTTTGCCCAGTGTCAGAAGAATCCAAACAGGTATT 539

Qy   1412 ggcttagttgtaagggctttaggatcagg 1440
          |||||| |||||||||| |||||||| |
Db    540 GGCTTANTTGTAAGGGCTNTAGGATCAAG 568
```

RESULT 3
AI870335/c
LOCUS AI870335 529 bp mRNA EST 07-MAR-2000
DEFINITION wl72f12.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2430479 3', mRNA sequence.
ACCESSION AI870335
VERSION AI870335.1 GI:5544303
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 529)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 567 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 455.
FEATURES Location/Qualifiers
source 1. .529
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2430479"
/clone_lib="NCI_CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 149 a 129 c 112 g 139 t

ORIGIN

Query Match 32.1%; Score 527; DB 26; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.9e-119;
Matches 527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1100 cagcctgtcaccgatggctgactgatgaaatggacacgtctcatctgacccactcttcct 1159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 CAGCCTGTCACCGATGGCTGACTGATGAAATGGACACGTCTCATCTGACCCACTCTTCCT 470

Qy   1160 tccactgaaggtcttcacgggcctccaggtggaccaaagggatgcacaggcggctcgcat 1219
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 TCCACTGAAGGTCTTCACGGGCCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCAT 410

Qy   1220 gccccagggccagctaagagttccaaagatctcagatttggttttagtcatgaatacata 1279
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 GCCCCAGGGCCAGCTAAGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATA 350

Qy   1280 aacagtctcaaactcgcacaatttttttccccctttgaaagccactggggccaatttgtg 1339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 AACAGTCTCAAACTCGCACAATTTTTTTCCCCCTTTTGAAAGCCACTGGGGCCAATTTGTG 290

Qy   1340 gttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgtcagaagaatc 1399
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GTTAAGAGGTGGTGAGATAAGAAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATC 230

Qy   1400 caagcaggtattggcttagttgtaagggctttaggatcaggctgaatatgaggacaaagt 1459
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CAAGCAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATCAGGCTGAATATGAGGACAAAGT 170

Qy   1460 gggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcattctgccacg 1519
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 GGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACG 110

Qy   1520 agagctaggtccttgatcttttctttagattgaaagtctgtctctgaacacaattatttg 1579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 AGAGCTAGGTCCTTGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATTTG 50

Qy   1580 taaaagttagtagttctttttaaatcattaaaagaggcttgctgaa 1626
          |||||||||||||||||||||||||||||||||||||||||||||
Db     49 TAAAAGTTAGTAGTTCTTTTTAAATCATTAAAAGAGGCTTGCTGAA 3
```

RESULT 4
BE465289/c
LOCUS BE465289 522 bp mRNA EST 27-JUL-2000
DEFINITION hw14c08.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3182894 3', mRNA sequence.
ACCESSION BE465289
VERSION BE465289.1 GI:9511064
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 522)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco

```
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3182894"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
BASE COUNT      147 a    125 c    106 g    144 t
ORIGIN


  Query Match          31.5%;  Score 517.2;  DB 108;  Length 522;
  Best Local Similarity   99.4%;  Pred. No. 4.9e-117;
  Matches  519;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy    1116 gctgactgatgaaatggacacgtctcatctgacccactcttccttccactgaaggtcttc 1175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GCTGACTGATGAAATGGACACGTCTCATCTGACCCACTCTTCCTTCCACTGAAGGTCTTC 463

Qy    1176 acgggcctccaggtggaccaaagggatgcacaggcggctcgcatgccccagggccagcta 1235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 ACGGGCCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTA 403

Qy    1236 agagttccaaagatctcagatttggtttttagtcatgaatacataaacagtctcaaactcg 1295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 AGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCG 343

Qy    1296 cacaatttttttccccctttttgaaagccactggggccaatttgtggttaagaggtggtgag 1355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 CACAATTTTTTCCCCCTTTTGAAAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGTGAG 283

Qy    1356 ataagaagtggaacgtgacatctttgccagttgtcagaagaatccaagcaggtattggct 1415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 ATAAGAAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCT 223

Qy    1416 tagttgtaagggctttaggatcaggctgaatatgaggacaaagtgggccacgttagcatc 1475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 TAGTTGTAAGGGCTTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATC 163

Qy    1476 tgcagagatcaatctggaggcttctgtttctgcattctgccacgagagctaggtccttga 1535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 TGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGA 103

Qy    1536 tcttttctttagattgaaagtctgtctctgaacacaattatttgtaaaagttagtagttc 1595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 TCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATTTGTAAAAGTTAGTAGTTC 43

Qy    1596 ttttttaaatcattaaaagaggcttgctgaaaaaaaaaaaaa 1637
           ||||||||||||||||||||||||||||||  | |||||||
Db      42 TTTTTTAAATCATTAAAAGAGGCTTGCTGAAGGATAAAAAAA 1


RESULT   5
AW015888/c
LOCUS       AW015888      503 bp    mRNA            EST       10-SEP-1999
DEFINITION  UI-H-BI0-aao-d-11-0-UI.s1 NCI_CGAP_Sub1 Homo sapiens cDNA clone
            IMAGE:2709956 3', mRNA sequence.
ACCESSION   AW015888
VERSION     AW015888.1  GI:5864645
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            NCI-CGAP clone distribution information can be found through the
            I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .503
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2709956"
                     /clone_lib="NCI_CGAP_Sub1"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NCI_CGAP_Sub1 library is a subtracted library derived from
                     BI.  BI constitutes a mixture of 21 normalized or
                     subtracted NCI_CGAP  libraries: NCI_CGAP_Co4,
                     NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10, NCI_CGAP_Co16
                     , NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3,
                     NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Co8,
                     NCI_CGAP_CLL1, NCI_CGAP_Lei2, NCI_CGAP_Brn23, NCI_CGAP_Lu5
                     , NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6
                     , NCI_CGAP_Brn25.  These 21 libraries were pooled and a
                     single-stranded DNA preparation of the resulting mixture
                     was used as a tracer in a subtractive hybridization with a
                     driver whose composition is detailed below: NCI_CGAP_Kid3
                     pool 1 LLAM 3334-3337, 3682-3683, 3798-3803   (IMAGE
                     CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855
                     ) NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725,
                     3776-3778   (IMAGE CloneIDs 1323912-1325831,
                     1471368-1472903, 1492104-1493255) NCI_CGAP_Lu5 pool 1 LLAM
                     3575-3582, 3851-3854   (IMAGE CloneIDs 1414920-1417991,
                     1520904-1522439)  NCI_CGAP_GC4 pool 1 LLAM 3164-3167,
                     3716-3720, 3733-3735   (IMAGE CloneIDs 1257096-1258631,
                     1469064-1470983, 1475592-1476743) NCI_CGAP_Pr22 pool 1
                     LLAM 2457-2459, 2758-2759, 3062-3068   (IMAGE CloneIDs
                     985608-986759, 1101192-1101959, 1217928-1220615)
                     NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872   (IMAGE
                     CloneIDs 1057416-1061255, 1144584-1145351)  The resulting
                     subtracted library contained 530,000 recombinants.
                     Subtraction was performed as previously described [Bonaldo
                     ,  Lennon & Soares (1996): Normalization and Subtraction:
                     Two Approaches  To Facilitate Gene Discovery. Genome
                     Research 6, 791-806.
                     TAG_LIB=NCI_CGAP_Lu5
                     TAG_TISSUE=lung
                     TAG_SEQ=CAAC"
BASE COUNT      139 a    118 c    101 g    145 t
ORIGIN


  Query Match          29.3%;  Score 480.6;  DB 38;  Length 503;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-108;
  Matches  486;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    1148 cccactcttccttccactgaaggtcttcacgggcctccaggtggaccaaagggatgcaca 1207
           |||||||||||||||| |||||||||||||||||||||||||| | ||||||||| |||
Db     498 CCCACTCTTCCTTCCAGTGAAGGTCTTCACGGGCCTCCAGGTGGCCAAAAGGGATGAACA 439
```

```
Qy   1208 ggcggctcgcatgccccagggccagctaagagttccaaagatctcagatttggtttttagt 1267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 GGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGTTCCAAAGATCTCAGATTTGGTTTTAGT 379

Qy   1268 catgaatacataaacagtctcaaactcgcacaatttttttccccctttttgaaagccactgg 1327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 CATGAATACATAAACAGTCTCAAACTCGCACAATTTTTTCCCCCTTTTGAAAGCCACTGG 319

Qy   1328 ggccaatttgtggttaagaggtggtgagataagaagtggaacgtgacatctttgccagtt 1387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 GGCCAATTTGTGGTTAAGAGGTGGTGAGATAAGAAGTGGAACGTGACATCTTTGCCAGTT 259

Qy   1388 gtcagaagaatccaagcaggtattggcttagttgtaagggctttaggatcaggctgaata 1447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 GTCAGAAGAATCCAAGCAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATCAGGCTGAATA 199

Qy   1448 tgaggacaaagtgggccacgttagcatctgcagagatcaatctggaggcttctgtttctg 1507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 TGAGGACAAAGTGGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTG 139

Qy   1508 cattctgccacgagagctaggtccttgatcttttctttagattgaaagtctgtctctgaa 1567
          ||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||||||
Db    138 CATTCTGCCACGAGAGCTAGGTCCTTGATCTTTTCTTTAGACTGAAAGTCTCTCTCTGAA 79

Qy   1568 cacaattatttgtaaaagttagtagttctttttttaaatcattaaaagaggcttgctgaaa 1627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     78 CACAATTATTTGTAAAAGTTAGTAGTTCTTTTTTAAATCATTAAAAGAGGCTTGCTGATG 19

Qy   1628 aaaaaaaaaaaaaaa 1642
           |||||||||||||||
Db     18 GAAAAAAAAAAAAAA 4
```

RESULT 6
AI817213/c
LOCUS AI817213 476 bp mRNA EST 07-MAR-2000
DEFINITION wg62h12.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2369735 3', mRNA sequence.
ACCESSION AI817213
VERSION AI817213.1 GI:5436281
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 476)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 574 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 461.
FEATURES Location/Qualifiers
source 1. .476
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2369735"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NbHSF pool 1: 309384-310919, 323208-325895 Soares Nb2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares Nb2HF8-9W pool 1: 758280-760583, 772104-774407 Soares NbHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NbHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 132 a 116 c 95 g 132 t 1 others
ORIGIN

Query Match 28.4%; Score 467; DB 25; Length 476;
Best Local Similarity 98.7%; Pred. No. 1.2e-104;
Matches 470; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy   1158 cttccactgaaggtcttcacgggcctccaggtggaccaaagggatgcacaggcggctcgc 1217
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    476 CTTCCACTGAAGGTCTTCACGGGCCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGC 417

Qy   1218 atgccccagggccagctaagagttccaaagatctcagatttggttttagtcatgaataca 1277
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    416 ATGCCCCAGGGCCAGCTAAGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACA 357

Qy   1278 taaacagtctcaaactcgcacaatttttttccccctttttgaaagccactggggccaatttg 1337
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    356 TAAACAGTCTCAAACTCGCACAATTTTTTCCCCCTTTTGAAAGCCACTGGGGCCAATTTG 297

Qy   1338 tggttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgtcagaagaa 1397
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    296 TGGTTAAGAGGTGGTGAGATAANAAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAA 237

Qy   1398 tccaagcaggtattggcttagttgtaagggctttaggatcaggctgaatatgaggacaaa 1457
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    236 TCCAAGCAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATCAGGCCGAATATGAGGACAAA 177

Qy   1458 gtgggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcattctgcca 1517
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    176 GTGGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCA 117

Qy   1518 cgagagctaggtccttgatcttttctttagattgaaagtctgtctctgaacacaattatt 1577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    116 CGAGAGCTAGGTCCTTGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATT 57

Qy   1578 tgtaaaagttagtagttctttttttaaatcattaaaagaggcttgctgaaaaaaaaa 1633
          |||||||||||| || |||||||||||||||||||||||||||||||||  |||||
Db     56 TGTAAAAGTTAGAAGCTCTTTTTTAAATCATTAAAAGAGGCTTGCTGAAGGAAAAA 1
```

RESULT 7
AA778133/c
LOCUS AA778133 461 bp mRNA EST 05-FEB-1998
DEFINITION zf46h06.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:380027 3', mRNA sequence.
ACCESSION AA778133
VERSION AA778133.1 GI:2837534
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 461)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

```
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 454.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Homo sapiens"
                     /db_xref="GDB:1288283"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:380027"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
BASE COUNT      132 a    114 c     88 g    127 t
ORIGIN


  Query Match             27.8%;  Score 456.2;  DB 11;  Length 461;
  Best Local Similarity   99.3%;  Pred. No. 5.3e-102;
  Matches  458;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1173 ttcacgggcctccaggtggaccaaagggatgcacaggcggctcgcatgccccagggccag 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 TTCACGGGCCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAG 402

Qy   1233 ctaagagttccaaagatctcagatttggttttagtcatgaatacataaacagtctcaaac 1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 CTAAGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAAC 342

Qy   1293 tcgcacaatttttttccccctttttgaaagccactggggccaatttgtggttaagaggtggt 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 TCGCACAATTTTTTCCCCCTTTTGAAAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGT 282

Qy   1353 gagataagaagtggaacgtgacatctttgccagttgtcagaagaatccaagcaggtattg 1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 GAGATAAGAAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTG 222

Qy   1413 gcttagttgtaagggctttaggatcaggctgaatatgaggacaaagtgggccacgttagc 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 GCTTAGTTGTAAGGGCTTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGC 162

Qy   1473 atctgcagagatcaatctggaggcttctgtttctgcattctgccacgagagctaggtcct 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 ATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCT 102

Qy   1533 tgatcttttctttagattgaaagtctgtctctgaacacaattatttgtaaaagttagtag 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 TGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATTTGTAAAAGTTAGTAG 42

Qy   1593 ttctttttaaatcattaaaagaggcttgctgaaaaaaaaa 1633
          ||||||||||||||||||||||||||||||||||  | |||
Db     41 TTCTTTTTTAAATCATTAAAAGAGGCTTGCTGAAGGATAAA 1


RESULT   8
AI492261/c
LOCUS       AI492261      487 bp    mRNA            EST       30-MAR-1999
DEFINITION  tf59b03.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2103533 3',
            mRNA sequence.
ACCESSION   AI492261
VERSION     AI492261.1  GI:4393264
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1155   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 359.
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2103533"
                     /clone_lib="NCI_CGAP_Brn23"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      132 a    113 c     90 g    152 t
ORIGIN


  Query Match             27.5%;  Score 452;  DB 21;  Length 487;
  Best Local Similarity   97.9%;  Pred. No. 5.8e-101;
  Matches  458;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1175 cacgggcctccaggtggaccaaagggatgcacaggcggctcgcatgccccagggccagct 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    487 CACGGGCCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCT 428

Qy   1235 aagagttccaaagatctcagatttggttttagtcatgaatacataaacagtctcaaactc 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    427 AAGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTC 368

Qy   1295 gcacaatttttttccccctttttgaaagccactggggccaatttgtggttaagaggtggtga 1354
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    367 GCACAATTTTTTCCCCCTTTTGAAAGCCCCTGGGGCCAATTTGTGGTTAAGAGGTGGTGA 308

Qy   1355 gataagaagtggaacgtgacatctttgccagttgtcagaagaatccaagcaggtattggc 1414
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 GATAAAAAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGC 248

Qy   1415 ttagttgtaagggctttaggatcaggctgaatatgaggacaaagtgggccacgttagcat 1474
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    247 TTAGTTGTAAGGGCTTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCCCGTTAGCAT 188
```

```
Qy   1475 ctgcagagatcaatctggaggcttctgtttctgcattctgccacgagagctaggtccttg 1534
          |||||||||||||| |||||||||||||| |||||||||| ||||||||||||||||||
Db    187 CTGCAGAGATCAATTTGGAGGCTTCTGTTTTTGCATTCTGCCCCGAGAGCTAGGTCCTTG 128

Qy   1535 atctttttctttagattgaaagtctgtctctgaacacaattatttgtaaaagttagtagtt 1594
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    127 ATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACCCAATTATTTGTAAAAGTTAGTAGTT 68

Qy   1595 ctttttaaatcattaaaagaggcttgctgaaaaaaaaaaaaaaaaa 1642
          |||||||||||||||||||||||||||||||  | ||||||||||||
Db     67 CTTTTTTAAATCATTAAAAGAGGCTTGCTGAAGGATAAAAAAAAAAA 20


RESULT  9
N52801/c
LOCUS       N52801        466 bp    mRNA            EST       15-FEB-1996
DEFINITION  yz18g05.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA
            clone IMAGE:283448 3', mRNA sequence.
ACCESSION   N52801
VERSION     N52801.1  GI:1193967
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 466)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 467.
FEATURES             Location/Qualifiers
     source          1. .466
                     /organism="Homo sapiens"
                     /db_xref="GDB:3901976"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:283448"
                     /clone_lib="Soares_multiple_sclerosis_2NbHMSP"
                     /sex="male"
                     /tissue_type="multiple sclerosis lesions"
                     /dev_stage="Age 46"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker V_TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
                     ; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis
                     lesions from one patient was kindly provided by Dr. Kevin
                     G. Becker (NINDS/NIH). "
BASE COUNT      131 a    117 c     91 g    127 t
ORIGIN


  Query Match             27.3%;  Score 448.8;  DB 143;  Length 466;
  Best Local Similarity   99.4%;  Pred. No. 3.5e-100;
  Matches  461;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;
```

```
Qy   1164 ctgaaggtcttcacgggcctccaggtggaccaaagggatgcacaggcggctcgcatgccc 1223
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 CTGAAGGTCTTCACGGGCCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCC 407

Qy   1224 cagggccagctaagagttccaaagatctcagatttggttttagtcatgaatacataaaca 1283
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 CAGGGCCAGCTAAGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACA 347

Qy   1284 gtctcaaactcgcacaatttttttccccctttgaaagccactggggccaatttgtggtta 1343
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 GTCTCAAACTCGCACAATTTTTTCCCCCTTTGAAAGCCACTGGGGCCAATTTGTGGTTA 287

Qy   1344 agaggtggtgagataagaagtggaacgtgacatctttgccagttgtcagaagaatccaag 1403
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 AGAGGTGGTGAGATAAGAAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAG 227

Qy   1404 caggtattggcttagttgtaagggctttaggatcaggctgaatatgaggacaaagtgggc 1463
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    226 CAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATCAGGCCGAATATGAGGACAAAGTGGGC 167

Qy   1464 cac-gttagcatctgcagagatcaatctggaggcttctgtttctgcattctgccacgaga 1522
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 CACAGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGA 107

Qy   1523 gctaggtccttgatcttttctttagattgaaagtctgtctctgaacacaattatttgtaa 1582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 GCTAGGTCCTTGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATTTGTAA 47

Qy   1583 aagttagtagtttctttttaaatcattaaaagaggcttgctgaa 1626
          ||||||| |||||||||||||||||||||||||||||||||||
Db     46 AAGTTAGAAGTTCTTTTTTAAATCATTAAAAGAGGCTTGCTGAA 3


RESULT  10
AW614686/c
LOCUS       AW614686      453 bp    mRNA            EST       23-MAR-2000
DEFINITION  hh31b08.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2956695 3',
            mRNA sequence.
ACCESSION   AW614686
VERSION     AW614686.1  GI:7319872
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 453)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            Seq primer: -40UP from Gibco
            High quality sequence stop: 448.
FEATURES             Location/Qualifiers
     source          1. .453
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2956695"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
```

modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 129 a 111 c 83 g 130 t
ORIGIN

Query Match 27.3%; Score 448.2; DB 93; Length 453;
Best Local Similarity 99.3%; Pred. No. 4.9e-100;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy   1185 caggtggaccaaagggatgcacaggcggctcgcatgccccagggccagctaagagttcca 1244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    453 CAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGTTCCA 394

Qy   1245 aagatctcagatttggttttagtcatgaatacataaacagtctcaaactcgcacaatttt 1304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 AAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCGCACAATTTT 334

Qy   1305 ttccccctttttgaaagccactggggccaatttgtggttaagaggtggtgagataagaagt 1364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 TTCCCCCTTTTGAAAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGTGAGATAAGAAGT 274

Qy   1365 ggaacgtgacatctttgccagttgtcagaagaatccaagcaggtattggcttagttgtaa 1424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 GGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCTTAGTTGTAA 214

Qy   1425 gggctttaggatcaggctgaatatgaggacaaagtgggccacgttagcatctgcagagat 1484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 GGGCTTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATCTGCAGAGAT 154

Qy   1485 caatctggaggcttctgtttctgcattctgccacgagagctaggtccttgatcttttctt 1544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 CAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGATCTTTTCTT 94

Qy   1545 tagattgaaagtctgtctctgaacacaattatttgtaaaagttagtagttctttttaaa 1604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 TAGATTGAAAGTCTGTCTCTGAACACAATTATTTGTAAAAGTTAGTAGTTCTTTTTAAA 34

Qy   1605 tcattaaaagaggcttgctgaaaaaaaaaaaa 1637
          |||||||||||||||||||||  | |||||||
Db     33 TCATTAAAAGAGGCTTGCTGAAGGATAAAAAAA 1
```

RESULT 11
AI569688/c
LOCUS AI569688 446 bp mRNA EST 14-APR-1999
DEFINITION tn04e05.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2166656 3', mRNA sequence.
ACCESSION AI569688
VERSION AI569688.1 GI:4533062
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 446)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 565 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 444.
FEATURES Location/Qualifiers
source 1. .446
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2166656"
/clone_lib="NCI_CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer {5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT T 3'}; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 129 a 112 c 84 g 121 t
ORIGIN

Query Match 26.9%; Score 442.4; DB 22; Length 446;
Best Local Similarity 99.8%; Pred. No. 1.3e-98;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy   1183 tccaggtggaccaaagggatgcacaggcggctcgcatgccccagggccagctaagagttc 1242
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    446 TCCAGGTGGACCAAGGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGTTC 387

Qy   1243 caaagatctcagatttggttttagtcatgaatacataaacagtctcaaactcgcacaatt 1302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    386 CAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCGCACAATT 327

Qy   1303 ttttccccctttgaaagccactggggccaatttgtggttaagaggtggtgagataagaa 1362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    326 TTTTCCCCCTTTTGAAAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGTGAGATAAGAA 267

Qy   1363 gtggaacgtgacatctttgccagttgtcagaagaatccaagcaggtattggcttagttgt 1422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    266 GTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCTTAGTTGT 207

Qy   1423 aagggctttaggatcaggctgaatatgaggacaaagtgggccacgttagcatctgcagag 1482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    206 AAGGGCTTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATCTGCAGAG 147

Qy   1483 atcaatctggaggcttctgtttctgcattctgccacgagagctaggtccttgatcttttc 1542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    146 ATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGATCTTTTC 87

Qy   1543 tttagattgaaagtctgtctctgaacacaattatttgtaaaagttagtagttctttttta 1602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     86 TTTAGATTGAAAGTCTGTCTCTGAACACAATTATTTGTAAAAGTTAGTAGTTCTTTTTTA 27

Qy   1603 aatcattaaaagaggcttgctgaa 1626
          |||||||||||||||||||||||
Db     26 AATCATTAAAAGAGGCTTGCTGAA 3
```

RESULT 12
AI199159/c
LOCUS AI199159 443 bp mRNA EST 01-DEC-1998
DEFINITION qi46h07.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:1859581 3', mRNA sequence.
ACCESSION AI199159

```
VERSION     AI199159.1  GI:3751755
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 573   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 442.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1859581"
                     /clone_lib="NCI_CGAP_Brn25"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      127 a    107 c     80 g    129 t
ORIGIN
```

```
  Query Match            26.8%;  Score 439.8;  DB 17;  Length 443;
  Best Local Similarity  99.5%;  Pred. No. 5.8e-98;
  Matches  441;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   1198 gggatgcacaggcggctcgcatgccccagggccagctaagagttccaaagatctcagatt 1257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 GGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGTTCCAAAGATCTCAGATT 384

Qy   1258 tggttttagtcatgaatacataaacagtctcaaactcgcacaatttttttccccctttttga 1317
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 TGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCGCACAATTTTTTCCCCCTTTTGA 324

Qy   1318 aagccactggggccaatttgtggttaagaggtggtgagataagaagtggaacgtgacatc 1377
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 AAGCCACTGGGGCCAATTTGTGGTTAAGAGGTGGTGAGATAAGAAGTGGAACGTGACATC 264

Qy   1378 tttgccagttgtcagaagaatccaagcaggtattggcttagttgtaagggctttaggatc 1437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 TTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATC 204

Qy   1438 aggctgaatatgaggacaaagtgggccacgttagcatctgcagagatcaatctggaggct 1497
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 AGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCT 144

Qy   1498 tctgtttctgcattctgccacgagagctaggtccttgatcttttctttagattgaaagtc 1557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 TCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGATCTTTTCTTTAGATTGAAAGTC 84

Qy   1558 tgtctctgaacacaattatttgtaaaagttagtagttctttttttaaatcattaaaagagg 1617
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 TGTCTCTGAACACAATTATTTGTAAAAGTTAGTAGTTCTTTTTTAAATCATTAAAAGAGG 24

Qy   1618 cttgctgaaaaaaaaaaaaaaaa 1640
          ||||||||| ||||||||||||
Db     23 CTTGCTGAAGGAAAAAAAAAAAA 1
```

```
RESULT  13
N92873/c
LOCUS       N92873         475 bp    mRNA            EST       20-AUG-1996
DEFINITION  zb71d08.s1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
            IMAGE:309039 3', mRNA sequence.
ACCESSION   N92873
VERSION     N92873.1  GI:1265182
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 475)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 550   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 327.
FEATURES             Location/Qualifiers
     source          1. .475
                     /organism="Homo sapiens"
                     /db_xref="GDB:1252452"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:309039"
                     /clone_lib="Soares_fetal_lung_NbHL19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTT-3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. This library was constructed
                     from the same fetus as the fetal heart library, Soares
                     fetal heart NbHH19W."
BASE COUNT      135 a    115 c     92 g    129 t      4 others
ORIGIN
```

```
  Query Match            25.7%;  Score 421.8;  DB 143;  Length 475;
  Best Local Similarity  96.8%;  Pred. No. 1.6e-93;
  Matches  460;  Conservative    0;  Mismatches   11;  Indels    4;  Gaps    3;
```

```
Qy   1163 actgaaggtcttcacgggcctccaggtgga-ccaaagggatgcacaggcggctcg--cat 1219
          ||||||| |||||||| | ||||| |||| |||| ||||||||||||||||    |||
Db    475 ACTGAAGNTCTTCACGGCCNTCCAGNTGGACCCAAGGGGATGCACAGGCGGCTTCGCCAT 416

Qy   1220 gccccagggccagctaagagttccaaagatct-cagatttggttttagtcatgaatacat 1278
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    415 GCCCCAGGGCCAGCTAAGAGTTCCAAAGATCTCCAGATTTGGTTTTAGTCATGAATACAT 356

Qy   1279 aaacagtctcaaactcgcacaatttttttccccctttgaaagccactggggccaatttgt 1338
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    355 AAACAGTCTNAAACTCGCACAATTTTTTCCCCCTTTTGAAAGCCACTGGGGCCAATTTGT 296

Qy   1339 ggttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgtcagaagaat 1398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    295 GGTTAAGAGGTGGTGAGATAAGAAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAAT 236

Qy   1399 ccaagcaggtattggcttagttgtaagggctttaggatcaggctgaatatgaggacaaag 1458
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    235 CCAAGCAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATCAGGCTGAATATGAGGACAAAG 176

Qy   1459 tgggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcattctgccac 1518
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    175 TGGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCAC 116

Qy   1519 gagagctaggtccttgatcttttcttttagattgaaagtctgtctctgaacacaattattt 1578
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    115 GAGAGCTAGGTCCTTGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATTT 56

Qy   1579 gtaaaagttagtagttctttttttaaatcattaaaagaggcttgctgaaaaaaaaa 1633
          |||||||||||||||||||||||||||||||||||||||||||||||  | |||
Db     55 GTAAAAGTTAGTAGTTCTTTTTTAAATCATTAAAAGAGGCTTGCTGAAGGATAAA 1
```

RESULT 14
AI371173/c
LOCUS AI371173 425 bp mRNA EST 16-FEB-1999
DEFINITION ta09c11.x1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:2043572 3', mRNA sequence.
ACCESSION AI371173
VERSION AI371173.1 GI:4149926
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 425)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 558 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 423.
FEATURES Location/Qualifiers
source 1. .425
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2043572"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 127 a 101 c 77 g 120 t
ORIGIN

Query Match 25.6%; Score 420.2; DB 19; Length 425;
Best Local Similarity 99.3%; Pred. No. 3.9e-93;
Matches 422; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy   1211 ggctcgcatgccccagggccagctaagagttccaaagatctcagatttggttttagtcat 1270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    425 GGCTCGCATGCCCCAGGGCCAGCTAAGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCAT 366

Qy   1271 gaatacataaacagtctcaaactcgcacaatttttttccccctttgaaagccactggggc 1330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    365 GAATACATAAACAGTCTCAAACTCGCACAATTTTTTCCCCCTTTTGAAAGCCACTGGGGC 306

Qy   1331 caatttgtggttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgtc 1390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    305 CAATTTGTGGTTAAGAGGTGGTGAGATAAGAAGTGGAACGTGACATCTTTGCCAGTTGTC 246

Qy   1391 agaagaatccaagcaggtattggcttagttgtaagggctttaggatcaggctgaatatga 1450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245 AGAAGAATCCAAGCAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATCAGGCTGAATATGA 186

Qy   1451 ggacaaagtgggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcat 1510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185 GGACAAAGTGGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCAT 126

Qy   1511 tctgccacgagagctaggtccttgatcttttcttttagattgaaagtctgtctctgaacac 1570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    125 TCTGCCACGAGAGCTAGGTCCTTGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACAC 66

Qy   1571 aattatttgtaaaagttagtagttctttttttaaatcattaaaagaggcttgctgaaaaaa 1630
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db     65 AATTATTTGTAAAAGTTAGTAGTTCTTTTTTAAATCATTAAAAGAGGCTTGCTGAAGGAT 6

Qy   1631 aaaaa 1635
          |||||
Db      5 AAAAA 1
```

RESULT 15
AI085287/c
LOCUS AI085287 419 bp mRNA EST 17-AUG-1998
DEFINITION qf18f01.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:1750393 3', mRNA sequence.
ACCESSION AI085287
VERSION AI085287.1 GI:3423710
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 419)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 392.

```
FEATURES             Location/Qualifiers
     source          1. .419
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1750393"
                     /clone_lib="NCI_CGAP_Brn25"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCCGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      126 a    101 c     78 g    114 t
ORIGIN
```

```
  Query Match             25.4%;  Score 417;  DB 16;  Length 419;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-92;
  Matches  417;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1210 cggctcgcatgccccagggccagctaagagttccaaagatctcagatttggttttagtca 1269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 CGGCTCGCATGCCCCAGGGCCAGCTAAGAGTTCCAAAGATCTCAGATTTGGTTTTAGTCA 360

Qy   1270 tgaatacataaacagtctcaaactcgcacaatttttttcccccttttgaaagccactgggg 1329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 TGAATACATAAACAGTCTCAAACTCGCACAATTTTTTCCCCCTTTTGAAAGCCACTGGGG 300

Qy   1330 ccaatttgtggttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgt 1389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 CCAATTTGTGGTTAAGAGGTGGTGAGATAAGAAGTGGAACGTGACATCTTTGCCAGTTGT 240

Qy   1390 cagaagaatccaagcaggtattggcttagttgtaagggctttaggatcaggctgaatatg 1449
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 CAGAAGAATCCAAGCAGGTATTGGCTTAGTTGTAAGGGCTTTAGGATCAGGCTGAATATG 180

Qy   1450 aggacaaagtgggccacgttagcatctgcagagatcaatctggaggcttctgtttctgca 1509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    179 AGGACAAAGTGGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTTCTGCA 120

Qy   1510 ttctgccacgagagctaggtccttgatcttttctttagattgaaagtctgtctctgaaca 1569
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119 TTCTGCCACGAGAGCTAGGTCCTTGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACA 60

Qy   1570 caattatttgtaaaagttagtagttctttttaaatcattaaaagaggcttgctgaa 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     59 CAATTATTTGTAAAAGTTAGTAGTTCTTTTTTAAATCATTAAAAGAGGCTTGCTGAA 3
```

Search completed: February 24, 2001, 21:04:33
Job time: 6986 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 21:06:11 ; Search time 86.16 Seconds
(without alignments)
3071.323 Million cell updates/sec

Title: US-09-320-713-3
Perfect score: 1642
Sequence: 1 ggaattcggcacgagctcgt.....tgaaaaaaaaaaaaaaaaaa 1642

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	56	3.4	3624	1	US-07-951-715A-6	Sequence 6, Appli
	2	56	3.4	3624	2	US-08-459-448A-6	Sequence 6, Appli
	3	56	3.4	3624	3	US-08-459-595A-6	Sequence 6, Appli
	4	56	3.4	3624	3	US-08-459-504B-6	Sequence 6, Appli
	5	56	3.4	3624	3	US-08-459-444-6	Sequence 6, Appli
	6	56	3.4	3624	3	US-09-053-549-7	Sequence 7, Appli
	7	53.4	3.3	2734	3	US-09-135-021-79	Sequence 79, Appl
	8	53.4	3.3	3181	3	US-09-135-021-1	Sequence 1, Appli
	9	53.4	3.3	3182	3	US-09-135-021-5	Sequence 5, Appli
	10	52.6	3.2	1485	1	US-08-471-601-23	Sequence 23, Appl
	11	52.6	3.2	1485	1	US-08-474-556-23	Sequence 23, Appl
	12	52.6	3.2	1485	1	US-08-479-382-23	Sequence 23, Appl
	13	52.6	3.2	1485	1	US-08-470-354-23	Sequence 23, Appl
	14	52.6	3.2	1485	1	US-08-479-383-23	Sequence 23, Appl
	15	52.6	3.2	1485	2	US-08-479-041-23	Sequence 23, Appl
	16	52	3.2	152331	3	US-09-128-155-16	Sequence 16, Appl
	17	51.4	3.1	1821	2	US-08-265-310-7	Sequence 7, Appli
	18	51.4	3.1	1821	3	US-08-951-742-7	Sequence 7, Appli
	19	50.8	3.1	2712	3	US-09-025-691-4	Sequence 4, Appli
C	20	50.6	3.1	8438	1	US-07-945-283-1	Sequence 1, Appli
	21	50.4	3.1	1777	1	US-08-173-508-5	Sequence 5, Appli
	22	50.4	3.1	1777	2	US-08-265-310-5	Sequence 5, Appli
	23	50.4	3.1	1777	3	US-08-951-742-5	Sequence 5, Appli
	24	50.4	3.1	2418	1	US-08-462-484-1	Sequence 1, Appli
	25	50.4	3.1	2418	1	US-08-441-147-1	Sequence 1, Appli
	26	50.4	3.1	2418	4	PCT-US95-07536-1	Sequence 1, Appli
	27	49.6	3.0	624	2	US-08-426-599B-3	Sequence 3, Appli
	28	49	3.0	1820	1	US-08-173-508-7	Sequence 7, Appli
	29	48.8	3.0	8854	3	US-09-053-549-1	Sequence 1, Appli
	30	48.4	2.9	1531	1	US-08-449-986-1	Sequence 1, Appli
	31	48.4	2.9	1531	2	US-08-756-855-1	Sequence 1, Appli
	32	48.2	2.9	1187	1	US-08-440-856A-2	Sequence 2, Appli
	33	48.2	2.9	38506	3	US-09-320-878-19	Sequence 19, Appl
	34	48.2	2.9	44377	2	US-08-804-227C-7	Sequence 7, Appli
	35	48.2	2.9	44377	2	US-08-804-198-1	Sequence 1, Appli
	36	47.8	2.9	1052	2	US-08-403-852D-10	Sequence 10, Appl
	37	47.8	2.9	1052	3	US-08-510-646B-10	Sequence 10, Appl
	38	47.4	2.9	1288	1	US-08-440-856A-9	Sequence 9, Appli
	39	46.8	2.9	2943	1	US-08-042-747A-7	Sequence 7, Appli
	40	46.6	2.8	2064	1	US-08-343-428-1	Sequence 1, Appli
	41	46.2	2.8	1389	2	US-08-023-980B-3	Sequence 3, Appli
	42	46.2	2.8	1389	2	US-08-486-953A-3	Sequence 3, Appli
	43	45.6	2.8	774	3	US-08-956-307B-12	Sequence 12, Appl
C	44	45.6	2.8	778	3	US-08-956-307B-11	Sequence 11, Appl
	45	45.2	2.8	1931	2	US-09-130-114-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-07-951-715A-6
; Sequence 6, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
;   APPLICANT: Koziel, Michael G.
;   APPLICANT: Desai, Nalini M.
;   APPLICANT: Lewis, Kelly S.
;   APPLICANT: Kramer, Vance C.
;   APPLICANT: Warren, Gregory W.
;   APPLICANT: Evola, Stephen V.
;   APPLICANT: Crossland, Lyle D.
;   APPLICANT: Wright, Martha S.
;   APPLICANT: Merlin, Ellis J.
;   APPLICANT: Launis, Karen L.
;   APPLICANT: Rothstein, Steven J.
;   APPLICANT: Bowman, Cindy G.
;   APPLICANT: Dawson, John L.
;   APPLICANT: Dunder, Erik M.
;   APPLICANT: Pace, Gary M.
;   APPLICANT: Suttie, Janet L.
;   TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;   TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
;   NUMBER OF SEQUENCES: 94
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: CIBA-GEIGY Corporation
;     STREET: 7 Skyline Drive
;     CITY: Hawthorne
;     STATE: New York
;     COUNTRY: USA
;     ZIP: 10532
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30B
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/07/951,715A
;     FILING DATE: 25-SEP-1992
;     CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/772,027
;     FILING DATE: 04-OCT-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Spruill, W. Murray
;     REGISTRATION NUMBER: 32,943
;     REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (919)541-8615
;     TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Synthetic DNA"
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3621
;      OTHER INFORMATION:  /product= "Full-length, maize
;      OTHER INFORMATION:  optmized cryIB"
;      OTHER INFORMATION:  /note= "Disclosed in Figure 6."
US-07-951-715A-6


  Query Match             3.4%;  Score 56;  DB 1;  Length 3624;
  Best Local Similarity   48.4%;  Pred. No. 0.0008;
  Matches  186;  Conservative    0;  Mismatches  195;  Indels    3;  Gaps    1;

Qy      83 gagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctgcg 142
           || | ||  |   | |   | | ||  |  ||||  |   | ||    |  || |  |
Db    2286 GATCGGCGAGAGCGAGCTGAAGGCCTACACCCGCTACCAGCTGCGCGGCTACATCGAGGA 2345

Qy     143 cagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtacct 202
           ||||  |   |    | ||||   |||  |||| || |    |    |  ||  | || |
Db    2346 CAGCCAGGACCTGGAGATCTACCTGATCCGCTACAACGCCAAGCACGAGACCCTGGACGT 2405

Qy     203 gcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggacgt 262
           ||| |   || |  |||||||  |       || ||    | |   |||||       | ||
Db    2406 GCCCGGCACCGAGAGCCTGTGGCCCCTGAGCGTGGAGAGCCCCATCGGCCGCTGCGGCGA 2465

Qy     263 gcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcctg 322
           || |  |||| |||||||   || |  |   | |  || ||||  | ||| |  |    |
Db    2466 GCCCAACCGCTGCGCCCCCCACTTCGAGTGGAACCCCGACCTGGACTGCAGCTGCC---G 2522

Qy     323 cgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacctg 382
           || |||||     | || | ||  |     | || ||| | |   | | | |     |||
Db    2523 CGACGGCGAGAAGTGCGCCCACCACAGCCACCACTTCAGCCTGGACATCGACGTGGGCTG 2582

Qy     383 cgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgc 442
           |  |  |  ||  |||||       |   |  |    | ||  ||| | |  |||| |
Db    2583 CACCGACCTGCACGAGAACCTGGGCGTGTGGGTGGTGTTCAAGATCAAGACCCAGGAGGG 2642

Qy     443 caagctcctgctgggccccaacga 466
           | |   ||  |||||| |   ||
Db    2643 CCACGCCCGCCTGGGCAACCTGGA 2666


RESULT   2
US-08-459-448A-6
; Sequence 6, Application US/08459448A
; Patent No. 5859336
;  GENERAL INFORMATION:
;    APPLICANT:  Koziel, Michael G.
;    APPLICANT:  Desai, Nalini M.
;    APPLICANT:  Lewis, Kelly S.
;    APPLICANT:  Kramer, Vance C.
;    APPLICANT:  Warren, Gregory W.
;    APPLICANT:  Evola, Stephen V.
;    APPLICANT:  Crossland, Lyle D.
;    APPLICANT:  Wright, Martha S.
;    APPLICANT:  Merlin, Ellis J.
;    APPLICANT:  Launis, Karen L.
;    APPLICANT:  Rothstein, Steven J.
;    APPLICANT:  Bowman, Cindy G.
;    APPLICANT:  Dawson, John L.
;    APPLICANT:  Dunder, Erik M.
;    APPLICANT:  Pace, Gary M.
;    APPLICANT:  Suttie, Janet L.
;    TITLE OF INVENTION:  SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY IN MAIZE
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5859336artis Corporation
;      STREET:  Patent & Trademark Dept., 520 White Plains
;      STREET:  Rd., POB 2005
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,448A
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/951,715
;      FILING DATE:  25-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/772,027
;      FILING DATE:  04-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  40403
;      REFERENCE/DOCKET NUMBER:  CGC 1577/CIP/DIV4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919)541-8582
;      TELEFAX:  (919)541-8689
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Synthetic DNA"
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3621
;      OTHER INFORMATION:  /product= "Full-length, maize
;      OTHER INFORMATION:  optmized cryIB"
;      OTHER INFORMATION:  /note= "Disclosed in Figure 6."
US-08-459-448A-6


  Query Match             3.4%;  Score 56;  DB 2;  Length 3624;
  Best Local Similarity   48.4%;  Pred. No. 0.0008;
  Matches  186;  Conservative    0;  Mismatches  195;  Indels    3;  Gaps    1;

Qy      83 gagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctgcg 142
           || | ||  |   | |   | | ||  |  ||||  |   | ||    |  || |  |
Db    2286 GATCGGCGAGAGCGAGCTGAAGGCCTACACCCGCTACCAGCTGCGCGGCTACATCGAGGA 2345

Qy     143 cagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtacct 202
           ||||  |   |    | ||||   |||  |||| || |    |    |  ||  | || |
Db    2346 CAGCCAGGACCTGGAGATCTACCTGATCCGCTACAACGCCAAGCACGAGACCCTGGACGT 2405

Qy     203 gcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggacgt 262
           ||| |   || |  |||||||  |       || ||    | |   |||||       | ||
Db    2406 GCCCGGCACCGAGAGCCTGTGGCCCCTGAGCGTGGAGAGCCCCATCGGCCGCTGCGGCGA 2465

Qy     263 gcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcctg 322
           || |  |||| |||||||   || |  |   | |  || ||||  | ||| |  |    |
Db    2466 GCCCAACCGCTGCGCCCCCCACTTCGAGTGGAACCCCGACCTGGACTGCAGCTGCC---G 2522

Qy     323 cgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacctg 382
           || |||||     | || | ||  |     | || ||| | |   | | | |     |||
Db    2523 CGACGGCGAGAAGTGCGCCCACCACAGCCACCACTTCAGCCTGGACATCGACGTGGGCTG 2582
```

```
Qy    383 cgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgc 442
          |  |  |  ||  |||||    |  |   |  |    | ||  ||| | |  ||||  |
Db   2583 CACCGACCTGCACGAGAACCTGGGCGTGTGGGTGGTGTTCAAGATCAAGACCCAGGAGGG 2642

Qy    443 caagctcctgctgggccccaacga 466
          | |   ||  ||||||  |   ||
Db   2643 CCACGCCCGCCTGGGCAACCTGGA 2666


RESULT   3
US-08-459-595A-6
; Sequence 6, Application US/08459595A
; Patent No. 6018104
;  GENERAL INFORMATION:
;    APPLICANT:  Koziel, Michael G.
;    APPLICANT:  Desai, Nalini M.
;    APPLICANT:  Lewis, Kelly S.
;    APPLICANT:  Kramer, Vance C.
;    APPLICANT:  Warren, Gregory W.
;    APPLICANT:  Evola, Stephen V.
;    APPLICANT:  Crossland, Lyle D.
;    APPLICANT:  Wright, Martha S.
;    APPLICANT:  Merlin, Ellis J.
;    APPLICANT:  Launis, Karen L.
;    APPLICANT:  Rothstein, Steven J.
;    APPLICANT:  Bowman, Cindy G.
;    APPLICANT:  Dawson, John L.
;    APPLICANT:  Dunder, Erik M.
;    APPLICANT:  Pace, Gary M.
;    APPLICANT:  Suttie, Janet L.
;    TITLE OF INVENTION:  SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY IN MAIZE
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018104artis Corporation
;      STREET:  Patent & Trademark Dept., 520 White Plains
;      STREET:  Rd., POB 2005
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,595A
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/951,715
;      FILING DATE:  25-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/772,027
;      FILING DATE:  04-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  40403
;      REFERENCE/DOCKET NUMBER:  CGC 1577/CIP/DIV3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919)541-8582
;      TELEFAX:  (919)541-8689
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Synthetic DNA"
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3621
;      OTHER INFORMATION:  /product= "Full-length, maize
;      OTHER INFORMATION:  optmized cryIB"
;      OTHER INFORMATION:  /note= "Disclosed in Figure 6."
US-08-459-595A-6


  Query Match             3.4%;  Score 56;  DB 3;  Length 3624;
  Best Local Similarity   48.4%;  Pred. No. 0.0008;
  Matches  186;  Conservative    0;  Mismatches  195;  Indels    3;  Gaps    1;

Qy     83 gagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctgcg 142
          || | ||  |   | |   | | ||  |  ||||  |   | ||    |  || | |
Db   2286 GATCGGCGAGAGCGAGCTGAAGGCCTACACCCGCTACCAGCTGCGCGGCTACATCGAGGA 2345

Qy    143 cagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtacct 202
          ||||  |   |    |  ||||   |||  ||||  || |  |    |  ||  | || |
Db   2346 CAGCCAGGACCTGGAGATCTACCTGATCCGCTACAACGCCAAGCACGAGACCCTGGACGT 2405

Qy    203 gcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggacgt 262
          |||  |   ||  |  |||||||  |       ||  ||    |  |   |||||      | ||
Db   2406 GCCCGGCACCGAGAGCCTGTGGCCCCTGAGCGTGGAGAGCCCCATCGGCCGCTGCGGCGA 2465

Qy    263 gcgcttccgcagcgccccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcctg 322
          ||  |  ||||  |||||||   ||  |  |   |  |  ||  ||||  |  |||  |  |   |
Db   2466 GCCCAACCGCTGCGCCCCCCACTTCGAGTGGAACCCCGACCTGGACTGCAGCTGCC---G 2522

Qy    323 cgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacctg 382
          ||  |||||     |  ||  |  ||   |  | ||  |||  |  |   |  |  |    |||
Db   2523 CGACGGCGAGAAGTGCGCCCACCACAGCCACCACTTCAGCCTGGACATCGACGTGGGCTG 2582

Qy    383 cgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgc 442
          |  |  |  ||  |||||    |  |   |  |    | ||  ||| | |  ||||  |
Db   2583 CACCGACCTGCACGAGAACCTGGGCGTGTGGGTGGTGTTCAAGATCAAGACCCAGGAGGG 2642

Qy    443 caagctcctgctgggccccaacga 466
          | |   ||  ||||||  |   ||
Db   2643 CCACGCCCGCCTGGGCAACCTGGA 2666


RESULT   4
US-08-459-504B-6
; Sequence 6, Application US/08459504B
; Patent No. 6075185
;  GENERAL INFORMATION:
;    APPLICANT:  Koziel, Michael G.
;    APPLICANT:  Desai, Nalini M.
;    APPLICANT:  Lewis, Kelly S.
;    APPLICANT:  Kramer, Vance C.
;    APPLICANT:  Warren, Gregory W.
;    APPLICANT:  Evola, Stephen V.
;    APPLICANT:  Crossland, Lyle D.
;    APPLICANT:  Wright, Martha S.
;    APPLICANT:  Merlin, Ellis J.
;    APPLICANT:  Launis, Karen L.
;    APPLICANT:  Rothstein, Steven J.
;    APPLICANT:  Bowman, Cindy G.
;    APPLICANT:  Dawson, John L.
;    APPLICANT:  Dunder, Erik M.
;    APPLICANT:  Pace, Gary M.
;    APPLICANT:  Suttie, Janet L.
;    TITLE OF INVENTION:  SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY IN MAIZE
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6075185artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/459,504B
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/459,595
;       FILING DATE:  02-JUN-1995
;       APPLICATION NUMBER:  US 07/951,715
;       FILING DATE:  25-SEP-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/772,027
;       FILING DATE:  04-OCT-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meigs, J. Timothy
;       REGISTRATION NUMBER:  38,241
;       REFERENCE/DOCKET NUMBER:  CGC1577/CIP/DIV
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (919)541-8587
;       TELEFAX:  (919)541-8689
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3624 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  other nucleic acid
;       DESCRIPTION:  /desc = "Synthetic DNA"
;     HYPOTHETICAL:  NO
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..3621
;       OTHER INFORMATION:  /product= "Full-length, maize
;       OTHER INFORMATION:  optmized cryIB"
;       OTHER INFORMATION:  /note= "Disclosed in Figure 6."
US-08-459-504B-6


  Query Match             3.4%;  Score 56;  DB 3;  Length 3624;
  Best Local Similarity   48.4%;  Pred. No. 0.0008;
  Matches  186;  Conservative    0;  Mismatches  195;  Indels    3;  Gaps    1;

Qy      83 gagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcccaccaacctgcg 142
           || | ||  |   | |  | | ||  |  ||||  |   | ||    |  || |  |
Db    2286 GATCGGCGAGAGCGAGCTGAAGGCCTACACCCGCTACCAGCTGCGCGGCTACATCGAGGA 2345

Qy     143 cagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtacct 202
           ||||  |   |    |  ||||   |||  |||| || |   |    |  ||  | || |
Db    2346 CAGCCAGGACCTGGAGATCTACCTGATCCGCTACAACGCCAAGCACGAGACCCTGGACGT 2405

Qy     203 gcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggacgt 262
           ||| |   || |  |||||||         || ||    | |   |||||      | ||
Db    2406 GCCCGGCACCGAGAGCCTGTGGCCCCTGAGCGTGGAGAGCCCCATCGGCCGCTGCGGCGA 2465

Qy     263 gcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcctg 322
           || |  |||| |||||||   || |  |   | |  || ||||  | ||| |  |    |
Db    2466 GCCCAACCGCTGCGCCCCCCACTTCGAGTGGAACCCCGACCTGGACTGCAGCTGCC---G 2522

Qy     323 cgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacctg 382
           || |||||     | || | ||  |     | || ||| | |   | | | |     |||
Db    2523 CGACGGCGAGAAGTGCGCCCACCACAGCCACCACTTCAGCCTGGACATCGACGTGGGCTG 2582

Qy     383 cgtccccgagccggagaaggacccagacagcatcaactccagcatcgacaaacagggcgc 442
           |  |  |  ||  |||||    |  |   |  |    | ||  ||| | |  ||||  |
Db    2583 CACCGACCTGCACGAGAACCTGCGCGTGTGGGTGGTGTTCAAGATCAAGACCCAGGAGGG 2642

Qy     443 caagctcctgctgggccccaacga 466
           | |   ||  ||||||  |   ||
Db    2643 CCACGCCCGCCTGGGCAACCTGGA 2666
```

```
RESULT   5
US-08-459-444-6
; Sequence 6, Application US/08459444A
; Patent No. 6121014
;   GENERAL INFORMATION:
;       APPLICANT: Koziel, Michael G.
;                  Desai, Nalini M.
;                  Lewis, Kelly S.
;                  Kramer, Vance C.
;                  Warren, Gregory W.
;                  Evola, Stephen V.
;                  Crossland, Lyle D.
;                  Wright, Martha S.
;                  Merlin, Ellis J.
;                  Launis, Karen L.
;       TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
;                           NUCLEIC ACID CODING SEQUENCE
;       NUMBER OF SEQUENCES: 94
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
;            STREET: 3054 Cornwallis Road
;            CITY: Research Triangle Park
;            STATE: NC
;            COUNTRY: USA
;            ZIP: 27709
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/459,444A
;            FILING DATE: 02-Jun-1995
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 07/951,715
;            FILING DATE: 25-SEP-1992
;            APPLICATION NUMBER: US 07/772,027
;            FILING DATE: 04-OCT-1991
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Meigs, J. Timothy
;            REGISTRATION NUMBER: 38,241
;            REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (919)541-8587
;            TELEFAX: (919)541-8689
;   INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 3624 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: other nucleic acid
;            DESCRIPTION: /desc = "Synthetic DNA"
;       HYPOTHETICAL: NO
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..3621
;            OTHER INFORMATION:  /product= "Full-length, maize
;            optmized cryIB"
;            /note= "Disclosed in Figure 6."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-459-444-6


  Query Match             3.4%;  Score 56;  DB 3;  Length 3624;
  Best Local Similarity   48.4%;  Pred. No. 0.0008;
  Matches  186;  Conservative    0;  Mismatches  195;  Indels    3;  Gaps    1;

Qy      83 gagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcccaccaacctgcg 142
           || | ||  |   | |  | | ||  |  ||||  |   | ||    |  || |  |
Db    2286 GATCGGCGAGAGCGAGCTGAAGGCCTACACCCGCTACCAGCTGCGCGGCTACATCGAGGA 2345
```

```
Qy    143 cagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtacct 202
          ||||  |   |    |  ||||   |||  |||| || |    |    |  ||  | || |
Db   2346 CAGCCAGGACCTGGAGATCTACCTGATCCGCTACAACGCCAAGCACGAGACCCTGGACGT 2405

Qy    203 gcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggacgt 262
          ||| |   || |  ||||||| |      || ||    | |   |||||      | ||
Db   2406 GCCCGGCACCGAGAGCCTGTGGCCCCTGAGCGTGGAGAGCCCCATCGGCCGCTGCGGCGA 2465

Qy    263 gcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcctg 322
          || |  |||| |||||||   || |  |   | |  || ||||  | ||| | |    |
Db   2466 GCCCAACCGCTGCGCCCCCCACTTCGAGTGGAACCCCGACCTGGACTGCAGCTGCC---G 2522

Qy    323 cgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacctg 382
          || |||||     | || | ||  |     | || ||| | |   | | | |     |||
Db   2523 CGACGGCGAGAAGTGCGCCCACCACAGCCACCACTTCAGCCTGGACATCGACGTGGGCTG 2582

Qy    383 cgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgc 442
          |  |  |  ||  ||||    |  |   |  |    | ||  ||| | |  |||| |
Db   2583 CACCGACCTGCACGAGAACCTGGGCGTGTGGGTGGTGTTCAAGATCAAGACCCAGGAGGG 2642

Qy    443 caagctcctgctgggccccaacga 466
          | |   ||  ||||||  |   ||
Db   2643 CCACGCCCGCCTGGGCAACCTGGA 2666
```

```
RESULT   6
US-09-053-549-7
; Sequence 7, Application US/09053549
; Patent No. 6121521
;  GENERAL INFORMATION:
;    APPLICANT:  Desai, Nalini
;    TITLE OF INVENTION:  No. 6121521el Insecticidal Protein and Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6121521artis Corporation
;      STREET:  3054 Cornwallis Rd.
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/053,549
;      FILING DATE:  01-APR-1998
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  40,403
;      REFERENCE/DOCKET NUMBER:  CGC 1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8582
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Synthetic DNA"
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3621
;      OTHER INFORMATION:  /product= "Full-length, maize
;      OTHER INFORMATION:  optmized cryIB"
;      OTHER INFORMATION:  /note= "Disclosed in Figure 6."
US-09-053-549-7
```

```
  Query Match             3.4%;  Score 56;  DB 3;  Length 3624;
  Best Local Similarity   48.4%;  Pred. No. 0.0008;
  Matches  186;  Conservative    0;  Mismatches  195;  Indels    3;  Gaps    1;

Qy     83 gagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctgcg 142
          || | ||  |   | |   | | ||  |  ||||  |   | ||    |  || |  |
Db   2286 GATCGGCGAGAGCGAGCTGAAGGCCTACACCCGCTACCAGCTGCGCGGCTACATCGAGGA 2345

Qy    143 cagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtacct 202
          ||||  |   |    |  ||||   |||  |||| || |    |    |  ||  | || |
Db   2346 CAGCCAGGACCTGGAGATCTACCTGATCCGCTACAACGCCAAGCACGAGACCCTGGACGT 2405

Qy    203 gcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggacgt 262
          ||| |   || |  ||||||| |      || ||    | |   |||||      | ||
Db   2406 GCCCGGCACCGAGAGCCTGTGGCCCCTGAGCGTGGAGAGCCCCATCGGCCGCTGCGGCGA 2465

Qy    263 gcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcctg 322
          || |  |||| |||||||   || |  |   | |  || ||||  | ||| | |    |
Db   2466 GCCCAACCGCTGCGCCCCCCACTTCGAGTGGAACCCCGACCTGGACTGCAGCTGCC---G 2522

Qy    323 cgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacctg 382
          || |||||     | || | ||  |     | || ||| | |   | | | |     |||
Db   2523 CGACGGCGAGAAGTGCGCCCACCACAGCCACCACTTCAGCCTGGACATCGACGTGGGCTG 2582

Qy    383 cgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgc 442
          |  |  |  ||  ||||    |  |   |  |    | ||  ||| | |  |||| |
Db   2583 CACCGACCTGCACGAGAACCTGGGCGTGTGGGTGGTGTTCAAGATCAAGACCCAGGAGGG 2642

Qy    443 caagctcctgctgggccccaacga 466
          | |   ||  ||||||  |   ||
Db   2643 CCACGCCCGCCTGGGCAACCTGGA 2666
```

```
RESULT   7
US-09-135-021-79
; Sequence 79, Application US/09135021A
; Patent No. 6150104
; GENERAL INFORMATION:
;  APPLICANT: Splawski, Igor
;  APPLICANT: Keating, Mark T.
;  TITLE OF INVENTION: A HOMOZYGOUS MUTATION IN KVLQT1 WHICH CAUSES JERVELL
;  TITLE OF INVENTION: AND LANGE-NIELSEN SYNDROME
;  FILE REFERENCE: 2323-128
;  CURRENT APPLICATION NUMBER: US/09/135,021A
;  CURRENT FILING DATE: 1998-08-17
;  EARLIER APPLICATION NUMBER: 08/874,655
;  EARLIER FILING DATE: 1997-06-13
;  EARLIER APPLICATION NUMBER: 60/094,477
;  EARLIER FILING DATE: 1998-07-29
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
;   LENGTH: 2734
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1743)
US-09-135-021-79
```

```
  Query Match             3.3%;  Score 53.4;  DB 3;  Length 2734;
  Best Local Similarity   47.2%;  Pred. No. 0.0028;
  Matches  162;  Conservative    0;  Mismatches  181;  Indels    0;  Gaps    0;

Qy    136 acctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacccca 195
          ||| |  ||  | |  |||||     | ||    |  ||  ||   |  | | | | |
Db   1406 accagtccattgggaagccctcactgttcatctccgtctcagaaaagagcaaggatcgcg 1465

Qy    196 ggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgagg 255
          |   |    |   | |  | |||||  ||| |        |  | |  |      ||
```

Thu Mar 1 14:24:21 2001

```
Db  1466 gcagcaacacgatcggcgcccgcctgaaccgagtagaagacaaggtgacgcagctggacc 1525
Qy   256 aggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcaccc 315
Db  1526 agaggctggcactcatcaccgacatgcttcaccagctgctctccttgcacggtggcagca 1585
Qy   316 ccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggct 375
Db  1586 ccccggcagcggcggcccccccagagagggcggggcccacatcacccagccctgcggca 1645
Qy   376 gcacctgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaac 435
Db  1646 gtggcggctccgtcgaccctgagctcttcctgcccagcaacaccctgcccacctacgagc 1705
Qy   436 agggcgccaagctcctgctgggccccaacgacgcgcccgctgg 478
Db  1706 agctgaccgtgcccaggaggggccccgatgaggggtcctgagg 1748

RESULT 8
US-09-135-021-1
; Sequence 1, Application US/09135021A
; Patent No. 6150104
; GENERAL INFORMATION:
;  APPLICANT: Splawski, Igor
;  APPLICANT: Keating, Mark T.
;  TITLE OF INVENTION: A HOMOZYGOUS MUTATION IN KVLQT1 WHICH CAUSES JERVELL
;  TITLE OF INVENTION: AND LANGE-NIELSEN SYNDROME
;  FILE REFERENCE: 2323-128
;  CURRENT APPLICATION NUMBER: US/09/135,021A
;  CURRENT FILING DATE: 1998-08-17
;  EARLIER APPLICATION NUMBER: 08/874,655
;  EARLIER FILING DATE: 1997-06-13
;  EARLIER APPLICATION NUMBER: 60/094,477
;  EARLIER FILING DATE: 1998-07-29
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;  LENGTH: 3181
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (163)..(2190)
US-09-135-021-1

 Query Match           3.3%;  Score 53.4;  DB 3;  Length 3181;
 Best Local Similarity  47.2%;  Pred. No. 0.003;
 Matches  162;  Conservative    0;  Mismatches  181;  Indels    0;  Gaps    0;

Qy   136 acctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacccca 195
Db  1853 accagtccattgggaagccctcactgttcatctccgtctcagaaaagagcaaggatcgcg 1912
Qy   196 ggtacctgcctgaagcctactgcctgtgccgggctgcctgaccgggctgttcggcgagg 255
Db  1913 gcagcaacacgatcggcgcccgcctgaaccgagtagaagacaaggtgacgcagctggacc 1972
Qy   256 aggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcaccc 315
Db  1973 agaggctggcactcatcaccgacatgcttcaccagctgctctccttgcacggtggcagca 2032
Qy   316 ccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggct 375
Db  2033 ccccggcagcggcggcccccccagagagggcggggcccacatcacccagccctgcggca 2092
Qy   376 gcacctgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaac 435
Db  2093 gtggcggctccgtcgaccctgagctcttcctgcccagcaacaccctgcccacctacgagc 2152
Qy   436 agggcgccaagctcctgctgggccccaacgacgcgcccgctgg 478
Db  2153 agctgaccgtgcccaggaggggccccgatgaggggtcctgagg 2195

RESULT 9
US-09-135-021-5
; Sequence 5, Application US/09135021A
; Patent No. 6150104
; GENERAL INFORMATION:
;  APPLICANT: Splawski, Igor
;  APPLICANT: Keating, Mark T.
;  TITLE OF INVENTION: A HOMOZYGOUS MUTATION IN KVLQT1 WHICH CAUSES JERVELL
;  TITLE OF INVENTION: AND LANGE-NIELSEN SYNDROME
;  FILE REFERENCE: 2323-128
;  CURRENT APPLICATION NUMBER: US/09/135,021A
;  CURRENT FILING DATE: 1998-08-17
;  EARLIER APPLICATION NUMBER: 08/874,655
;  EARLIER FILING DATE: 1997-06-13
;  EARLIER APPLICATION NUMBER: 60/094,477
;  EARLIER FILING DATE: 1998-07-29
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;  LENGTH: 3182
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (163)..(1011)
;  FEATURE:
;  NAME/KEY: mutation
;  LOCATION: (730)
;  OTHER INFORMATION: This base is an insertion as compared to the
;  OTHER INFORMATION: wild-type.
US-09-135-021-5

 Query Match           3.3%;  Score 53.4;  DB 3;  Length 3182;
 Best Local Similarity  47.2%;  Pred. No. 0.003;
 Matches  162;  Conservative    0;  Mismatches  181;  Indels    0;  Gaps    0;

Qy   136 acctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacccca 195
Db  1854 accagtccattgggaagccctcactgttcatctccgtctcagaaaagagcaaggatcgcg 1913
Qy   196 ggtacctgcctgaagcctactgcctgtgccgggctgcctgaccgggctgttcggcgagg 255
Db  1914 gcagcaacacgatcggcgcccgcctgaaccgagtagaagacaaggtgacgcagctggacc 1973
Qy   256 aggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcac 315
Db  1974 agaggctggcactcatcaccgacatgcttcaccagctgctctccttgcacggtgc 2033
Qy   316 ccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatc 375
Db  2034 ccccggcagcggcggcccccccagagagggcggggcccacatca 2093
Qy   376 gcacctgcgtccccgagccggagaaggacgcagacagcatc 435
Db  2094 gtggcggctccgtcgaccctgagctcttcctgcccag 2153
Qy   436 agggcgccaagctcctgctgggccccaacgacg 478
Db  2154 agctgaccgtgcccaggaggggccccgat 2195

RESULT 10
US-08-471-601-23
; Sequence 23, Application US
; Patent No. 5689049
;  GENERAL INFORMATION:
;  APPLICANT: CIGAN,
;  APPLICANT: ALBE
;  TITLE OF INVENT
;
```

...C C.
...eversible Nuclear Genetic System For

```
;    TITLE OF INVENTION:  Male Sterility In Transgenic Plants
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,601
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/351,899
;      FILING DATE:  08-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/341/PIHI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1485 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-471-601-23


  Query Match             3.2%;  Score 52.6;  DB 1;  Length 1485;
  Best Local Similarity   49.5%;  Pred. No. 0.0034;
  Matches  164;  Conservative    0;  Mismatches  164;  Indels    3;  Gaps    1;

Qy      71 ggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcc 130
           | || || | ||| ||| ||||| || | |  || |   |  |||    |  ||| |
Db     486 GTCGAGCGAGGCGCGCTTCCCGGGCGGCGACCTGCACCTGGCGCGCGCGCTGGGGCTGAG 545

Qy     131 caccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggta 190
           | |  || | ||| ||||   ||       |  | |   |||   || |  |||| ||
Db     546 CCCGGACGTCCGCCGCGTCATGCTGGCCTTCACCGGCTGCTCGGGCGGCGTGGCGGGGCT 605

Qy     191 ccccaggtacctgcctgaagccta---ctgcctgtgccggggctgcctgaccgggctgtt 247
           || |  |  |  |      ||| |   ||||| | ||  | ||   |||   |     |
Db     606 CCGCGTGGCCAAGGGCCTGGCCGAGAGCTGCCCGGGCGCGCGCGTGCTGCTGGCCACCTC 665

Qy     248 cggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcg 307
           ||          |||| | ||||||   ||     | ||   |||  || | || || |
Db     666 CGAGACCACCATCGTGGGGTTCCGCCCGCCCAGCCCCGACCGCCCCTACGACCTCGTGGG 725

Qy     308 ccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccc 367
           |    | | |  | ||| ||||   |  ||| |  || ||   || ||  | ||  | |
Db     726 CGTGGCGCTCTTCGGCGACGGCGCGGGCGCCGCCGTCATCGGCACCGACCCCGCCCCCGC 785

Qy     368 cgtgggctgcacctgcgtccccgagccggag 398
           || | ||     || ||  | | |  ||| |
Db     786 CGAGCGCCCGCTCTTCGAGCTCCACTCGGCG 816


RESULT  11
US-08-474-556-23
; Sequence 23, Application US/08474556
; Patent No. 5689051
;  GENERAL INFORMATION:
;    APPLICANT:  CIGAN, Andrew M.
;    APPLICANT:  ALBERTSEN, Marc C.
;    TITLE OF INVENTION:  Reversible Nuclear Genetic System For
;    TITLE OF INVENTION:  Male Sterility In Transgenic Plants
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,556
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/351,899
;      FILING DATE:  08-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/329/PIHI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1485 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-474-556-23


  Query Match             3.2%;  Score 52.6;  DB 1;  Length 1485;
  Best Local Similarity   49.5%;  Pred. No. 0.0034;
  Matches  164;  Conservative    0;  Mismatches  164;  Indels    3;  Gaps    1;

Qy      71 ggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcc 130
           | || || | ||| ||| ||||| || | |  || |   |  |||    |  ||| |
Db     486 GTCGAGCGAGGCGCGCTTCCCGGGCGGCGACCTGCACCTGGCGCGCGCGCTGGGGCTGAG 545

Qy     131 caccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggta 190
           | |  || | ||| ||||   ||       |  | |   |||   || |  |||| ||
Db     546 CCCGGACGTCCGCCGCGTCATGCTGGCCTTCACCGGCTGCTCGGGCGGCGTGGCGGGGCT 605

Qy     191 ccccaggtacctgcctgaagccta---ctgcctgtgccggggctgcctgaccgggctgtt 247
           || |  |  |  |      ||| |   ||||| | ||  | ||   |||   |     |
Db     606 CCGCGTGGCCAAGGGCCTGGCCGAGAGCTGCCCGGGCGCGCGCGTGCTGCTGGCCACCTC 665

Qy     248 cggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcg 307
           ||          |||| | ||||||   ||     | ||   |||  || | || || |
Db     666 CGAGACCACCATCGTGGGGTTCCGCCCGCCCAGCCCCGACCGCCCCTACGACCTCGTGGG 725

Qy     308 ccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccc 367
           |    | | |  | ||| ||||   |  ||| |  || ||   || ||  | ||  | |
Db     726 CGTGGCGCTCTTCGGCGACGGCGCGGGCGCCGCCGTCATCGGCACCGACCCCGCCCCCGC 785

Qy     368 cgtgggctgcacctgcgtccccgagccggag 398
           || | ||     || ||  | | |  ||| |
Db     786 CGAGCGCCCGCTCTTCGAGCTCCACTCGGCG 816


RESULT  12
US-08-479-382-23
; Sequence 23, Application US/08479382
; Patent No. 5763243
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  CIGAN, Andrew M.
;    APPLICANT:  ALBERTSEN, Marc C.
;    TITLE OF INVENTION:  Reversible Nuclear Genetic System For
;    TITLE OF INVENTION:  Male Sterility In Transgenic Plants
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,382
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/351,899
;      FILING DATE:  08-DEC-1994
;    ATTORNEY/AGENT INFORMATION
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/339/PIHI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1485 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-479-382-23


  Query Match             3.2%;  Score 52.6;  DB 1;  Length 1485;
  Best Local Similarity   49.5%;  Pred. No. 0.0034;
  Matches  164;  Conservative    0;  Mismatches  164;  Indels    3;  Gaps    1;

Qy      71 ggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcc 130
           | || || | ||| ||| ||||| || | |  || |   |  |||    |  ||| |
Db     486 GTCGAGCGAGGCGCGCTTCCCGGGCGGCGACCTGCACCTGGCGCGCGCGCTGGGGCTGAG 545

Qy     131 caccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggta 190
           | |  || | ||| ||||   |         |  | |   |||   || |  |||| ||
Db     546 CCCGGACGTCCGCCGCGTCATGCTGGCCTTCACCGGCTGCTCGGGCGGCGTGGCGGGGCT 605

Qy     191 ccccaggtacctgcctgaagccta---ctgcctgtgccgggggctgcctgaccgggctgtt 247
           || |  |  |  |      ||| |   ||||| | ||  | ||   |||   |     |
Db     606 CCGCGTGGCCAAGGGCCTGGCCGAGAGCTGCCCGGGCGCGCGCGTGCTGCTGGCCACCTC 665

Qy     248 cggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcg 307
           ||          |||| | ||||||    ||     | ||   |||  || | || || |
Db     666 CGAGACCACCATCGTGGGGTTCCGCCCGCCCAGCCCCGACCGCCCCTACGACCTCGTGGG 725

Qy     308 ccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccc 367
           |    | | |  | ||| |||||   |  ||| |  || ||   || ||  | ||  | |
Db     726 CGTGGCGCTCTTCGGCGACGGCGCGGGCGCCGCCGTCATCGGCACCGACCCCGCCCCCGC 785

Qy     368 cgtgggctgcacctgcgtccccgagccggag 398
           || | ||     || ||  | | |  ||| |
Db     786 CGAGCGCCCGCTCTTCGAGCTCCACTCGGCG 816


RESULT  13
US-08-470-354-23
```

```
; Sequence 23, Application US/08470354
; Patent No. 5792853
;  GENERAL INFORMATION:
;    APPLICANT:  CIGAN, Andrew M.
;    APPLICANT:  ALBERTSEN, Marc C.
;    TITLE OF INVENTION:  Reversible Nuclear Genetic System For
;    TITLE OF INVENTION:  Male Sterility In Transgenic Plants
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,354
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/351,899
;      FILING DATE:  08-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/337/PIHI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1485 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-470-354-23


  Query Match             3.2%;  Score 52.6;  DB 1;  Length 1485;
  Best Local Similarity   49.5%;  Pred. No. 0.0034;
  Matches  164;  Conservative    0;  Mismatches  164;  Indels    3;  Gaps    1;

Qy      71 ggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcc 130
           | || || | ||| ||| ||||| || | |  || |   |  |||    |  ||| |
Db     486 GTCGAGCGAGGCGCGCTTCCCGGGCGGCGACCTGCACCTGGCGCGCGCGCTGGGGCTGAG 545

Qy     131 caccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggta 190
           | |  || | ||| ||||   ||        |  | |   |||   || |  |||| ||
Db     546 CCCGGACGTCCGCCGCGTCATGCTGGCCTTCACCGGCTGCTCGGGCGGCGTGGCGGGGCT 605

Qy     191 ccccaggtacctgcctgaagcctа---ctgcctgtgccgggggctgcctgaccgggctgtt 247
           || |  |  |  |      ||| |   ||||| | ||  | ||   |||   |     |
Db     606 CCGCGTGGCCAAGGGCCTGGCCGAGAGCTGCCCGGGCGCGCGCGTGCTGCTGGCCACCTC 665

Qy     248 cggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcg 307
           ||          |||| | ||||||    ||     | ||   |||  || | || || |
Db     666 CGAGACCACCATCGTGGGGTTCCGCCCGCCCAGCCCCGACCGCCCCTACGACCTCGTGGG 725

Qy     308 ccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccc 367
           |    | | |  | ||| |||||   |  ||| |  || ||   || ||  | ||  | |
Db     726 CGTGGCGCTCTTCGGCGACGGCGCGGGCGCCGCCGTCATCGGCACCGACCCCGCCCCCGC 785

Qy     368 cgtgggctgcacctgcgtccccgagccggag 398
           || | ||     || ||  | | |  ||| |
Db     786 CGAGCGCCCGCTCTTCGAGCTCCACTCGGCG 816
```

```
RESULT 14
US-08-479-383-23
; Sequence 23, Application US/08479383
; Patent No. 5795753
; GENERAL INFORMATION:
;   APPLICANT: CIGAN, Andrew M.
;   APPLICANT: ALBERTSEN, Marc C.
;   TITLE OF INVENTION: Reversible Nuclear Genetic System For
;   TITLE OF INVENTION: Male Sterility In Transgenic Plants
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Foley & Lardner
;     STREET: 3000 K Street, N.W., Suite 500
;     CITY: Washington
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20007-5109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/479,383
;     FILING DATE: 07-JUN-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/351,899
;     FILING DATE: 08-DEC-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BENT, Stephen A.
;     REGISTRATION NUMBER: 29,768
;     REFERENCE/DOCKET NUMBER: 33229/340/PIHI
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (202)672-5300
;     TELEFAX: (202)672-5399
;     TELEX: 904136
; INFORMATION FOR SEQ ID NO: 23:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1485 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-479-383-23


  Query Match             3.2%;  Score 52.6;  DB 1;  Length 1485;
  Best Local Similarity   49.5%;  Pred. No. 0.0034;
  Matches  164;  Conservative    0;  Mismatches  164;  Indels    3;  Gaps    1;

Qy      71 ggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcc 130
           | || || | ||| ||| |||||  || | |  || |   |  |||    |  ||| |
Db     486 GTCGAGCGAGGCGCGCTTCCCGGGCGGCGACCTGCACCTGGCGCGCGCGCTGGGGCTGAG 545

Qy     131 caccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggta 190
           | |  || | ||| ||||   ||       |  | |   |||   || |  |||| ||
Db     546 CCCGGACGTCCGCCGCGTCATGCTGGCCTTCACCGGCTGCTCGGGCGGCGTGGCGGGGCT 605

Qy     191 ccccaggtacctgcctgaagccta---ctgcctgtgccggggctgcctgaccgggctgtt 247
           || |  |  |  |      ||| |   ||||| | ||  | ||   |||    |     |
Db     606 CCGCGTGGCCAAGGGCCTGGCCGAGAGCTGCCCGGGCGCGCGCGTGCTGCTGGCCACCTC 665

Qy     248 cggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcg 307
           ||          |||| | ||||||    ||     | ||   |||  || | || || |
Db     666 CGAGACCACCATCGTGGGGTTCCGCCCGCCCAGCCCCGACCGCCCCTACGACCTCGTGGG 725

Qy     308 ccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccc 367
           |    | | |  | ||| |||||   |  ||| |  || ||   || ||  | ||  | |
Db     726 CGTGGCGCTCTTCGGCGACGGCGCGGGCGCCGCCGTCATCGGCACCGACCCCGCCCCCGC 785

Qy     368 cgtgggctgcacctgcgtccccgagccggag 398
           || | ||     || ||  | | |  ||| |
Db     786 CGAGCGCCCGCTCTTCGAGCTCCACTCGGCG 816
```

```
RESULT 15
US-08-479-041-23
; Sequence 23, Application US/08479041
; Patent No. 5837851
; GENERAL INFORMATION:
;   APPLICANT: CIGAN, Andrew M.
;   APPLICANT: ALBERTSEN, Marc C.
;   TITLE OF INVENTION: Reversible Nuclear Genetic System For
;   TITLE OF INVENTION: Male Sterility In Transgenic Plants
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Foley & Lardner
;     STREET: 3000 K Street, N.W., Suite 500
;     CITY: Washington
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20007-5109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/479,041
;     FILING DATE: 07-JUN-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/351,899
;     FILING DATE: 08-DEC-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BENT, Stephen A.
;     REGISTRATION NUMBER: 29,768
;     REFERENCE/DOCKET NUMBER: 33229/338/PIHI
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (202)672-5300
;     TELEFAX: (202)672-5399
;     TELEX: 904136
; INFORMATION FOR SEQ ID NO: 23:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1485 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-479-041-23


  Query Match             3.2%;  Score 52.6;  DB 2;  Length 1485;
  Best Local Similarity   49.5%;  Pred. No. 0.0034;
  Matches  164;  Conservative    0;  Mismatches  164;  Indels    3;  Gaps    1;

Qy      71 ggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcc 130
           | || || | ||| ||| |||||  || | |  || |   |  |||    |  ||| |
Db     486 GTCGAGCGAGGCGCGCTTCCCGGGCGGCGACCTGCACCTGGCGCGCGCGCTGGGGCTGAG 545

Qy     131 caccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggta 190
           | |  || | ||| ||||   ||       |  | |   |||   || |  |||| ||
Db     546 CCCGGACGTCCGCCGCGTCATGCTGGCCTTCACCGGCTGCTCGGGCGGCGTGGCGGGGCT 605

Qy     191 ccccaggtacctgcctgaagccta---ctgcctgtgccggggctgcctgaccgggctgtt 247
           || |  |  |  |      ||| |   ||||| | ||  | ||   |||    |     |
Db     606 CCGCGTGGCCAAGGGCCTGGCCGAGAGCTGCCCGGGCGCGCGCGTGCTGCTGGCCACCTC 665

Qy     248 cggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcg 307
           ||          |||| | ||||||    ||     | ||   |||  || | || || |
Db     666 CGAGACCACCATCGTGGGGTTCCGCCCGCCCAGCCCCGACCGCCCCTACGACCTCGTGGG 725

Qy     308 ccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccc 367
           |    | | |  | ||| |||||   |  ||| |  || ||   || ||  | ||  | |
Db     726 CGTGGCGCTCTTCGGCGACGGCGCGGGCGCCGCCGTCATCGGCACCGACCCCGCCCCCGC 785

Qy     368 cgtgggctgcacctgcgtccccgagccggag 398
```

```
          || | ||     || ||  | | |  ||| |
Db    786 CGAGCGCCCGCTCTTCGAGCTCCACTCGGCG 816
```

Search completed: February 24, 2001, 21:06:35
Job time: 6993 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 21:51:30 ; Search time 119.29 Seconds
(without alignments)
5170.912 Million cell updates/sec

Title: US-09-320-713-3
Perfect score: 1642
Sequence: 1 ggaattcggcacgagctcgt.....tgaaaaaaaaaaaaaaaaaa 1642

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1642	100.0	1642	21	Z36835	Partial nucleotide
	2	1615.6	98.4	2361	21	Z52201	Human transforming
	3	1598.6	97.4	1819	21	Z52195	Human transforming
	4	463.4	28.2	522	21	Z36837	Partial nucleotide
	5	326.4	19.9	1221	21	Z52198	Murine transformin
	6	239.6	14.6	498	21	Z36853	cDNA clone HAGBX04
	7	196	11.9	371	21	Z36852	cDNA clone HE2CD08
c	8	167.4	10.2	178	21	Z36854	cDNA clone HCEBA24
	9	165.4	10.1	264	21	Z36855	cDNA clone HCELE59
	10	73.2	4.5	540	20	X89401	Human interleukin-
	11	73.2	4.5	687	20	Z34282	Human PRO1031 nucl
	12	73.2	4.5	687	21	Z29727	Human Interleukin
	13	73.2	4.5	705	20	X15870	cDNA encoding inte
	14	73.2	4.5	736	20	V83172	Human Zcyto7 cDNA.
	15	73.2	4.5	738	20	X84472	Human EDIRF I codi
	16	73.2	4.5	1259	20	X84474	Human EDIRF II cod
	17	71.8	4.4	100	16	T23250	Human gene signatu
c	18	61.6	3.8	114955	20	X53491	Human adenosine A1
	19	60.2	3.7	114955	20	X53491	Human adenosine A1
	20	59.4	3.6	692	20	V83179	Murine Zcyto7 cDNA
	21	59.4	3.6	756	20	X84473	Murine EDIRF I cod
	22	58.8	3.6	540	20	X89402	Mouse interleukin-
	23	56.8	3.5	591	21	A09153	Human interleukin-
	24	56.8	3.5	1047	21	Z29728	Human Interleukin
	25	56.8	3.5	1067	21	Z36836	Nucleotide sequenc
	26	56.8	3.5	4356	16	Q95540	Cardiac adenylyl c
	27	56.2	3.4	1145	16	T04785	Cellobiase gene fr
	28	55.2	3.4	2291	9	N80309	Entire amylase gen
	29	55.2	3.4	4356	14	Q37543	Cardiac adenylyl c
	30	53.6	3.3	265	20	X84482	Probe for murine E
	31	53.6	3.3	497	20	V83181	Murine Zcyto7 expr
c	32	53.6	3.3	1127	21	A02477	Human colon cancer
	33	53.6	3.3	2271	11	Q06844	Amylase gene from
	34	53.4	3.3	2821	18	T94004	DNA encoding human
	35	53.4	3.3	2821	18	T90730	Human KVLQT1 full-
	36	53.4	3.3	3181	21	Z90669	Human KVLQT1 prote
	37	53.4	3.3	3181	21	Z98901	Human long QT synd
c	38	53	3.2	1059	21	A01833	Human colon cancer
	39	53	3.2	24379	18	T93095	Streptomyces freno
	40	53	3.2	24379	19	V25925	Streptomyces roseo
	41	52.6	3.2	1485	17	T35134	5126 anther specif
	42	52.4	3.2	830	21	Z29729	Virtual DNA fragme
	43	52	3.2	3198	20	X02974	Human IL-1ra BAC c
	44	51.6	3.1	1497	11	Q03367	Phenol oxidase (PO
	45	51.6	3.1	1497	11	Q03571	Sequence encoding

ALIGNMENTS

RESULT 1
Z36835
ID Z36835 standard; cDNA; 1642 BP.
XX
AC Z36835;
XX
DT 13-MAR-2000 (first entry)
XX
DE Partial nucleotide sequence of human interleukin-22.
XX
KW Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW immune cell chemotaxis; haematopoietic cell disorder;
KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW inflammation; hyperproliferative disorder; tissue regeneration;
KW embryonic stem cell differentiation; embryonic stem cell proliferation;
KW haematopoietic lineage; allergic asthma; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 3..485
FT /*tag= a
FT /product= "partial IL-22"
XX
PN WO9961617-A1.
XX
PD 02-DEC-1999.
XX
PF 27-MAY-1999; 99WO-US11644.
XX
PR 29-MAY-1998; 98US-0087340.
PR 10-SEP-1998; 98US-0099805.
PR 30-APR-1999; 99US-0131965.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ebner R;
XX
DR WPI; 2000-072622/06.
DR P-PSDB; Y53891.
XX
PT Novel polynucleotides used to develop products for treating e.g. immune
PT disorders, blood disorders, autoimmune disorders, allergies,
PT inflammation, hyperproliferative disorders or infections -
XX
PS Claim 12; Fig 2A-B; 170pp; English.
XX
CC The present sequence encodes a partial human interleukin-22 (IL-22)
CC protein. The specification also describes IL-21 polynucleotides and
CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating
CC deficiencies or disorders of the immune system, by activating or
CC inhibiting the proliferation, differentiation, or mobilization
CC (chemotaxis) of immune cells, treating or detecting deficiencies
CC of haematopoietic cells, to modulate haemostatic or thrombolytic
CC activity, in treating or detecting autoimmune disorders, treating
CC asthma (particularly allergic asthma) or other respiratory problems,
CC to treat and/or prevent organ rejection or graft-versus-host disease
CC (GVHD), to modulate inflammation, to treat or detect hyperproliferative
CC disorders, to treat or detect infectious agents, to differentiate,
CC proliferate and attract cells, leading to the regeneration of tissues,
CC IL-21 and IL-22 may also increase or decrease the differentiation or
CC proliferation of embryonic stem cells and haematopoietic lineage, may
CC be used to modulate mammalian characteristics.
XX
SQ Sequence 1642 BP; 377 A; 458 C; 439 G; 368 T; 0 other;

Query Match 100.0%; Score 1642; DB 21; Length 1642;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ggaattcggcacgagctcgtgccgtgctcagtgccttccaccacacgctgcagctggggc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ggaattcggcacgagctcgtgccgtgctcagtgccttccaccacacgctgcagctggggc 60

Qy     61 cgcgtgagcaggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 cgcgtgagcaggcgcgcaacgcgagctgcccggcaggggcaggcccgccgaccgccgct 120

Qy    121 tccggccgcccaccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tccggccgcccaccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacc 180

Qy    181 cggcgaggtaccccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 cggcgaggtaccccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccg 240

Qy    241 ggctgttcggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ggctgttcggcgaggaggacgtgcgcttccgcagcgcccctgtctacatgcccaccgtcg 300

Qy    301 tcctgcgccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtca 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 tcctgcgccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtca 360

Qy    361 ccatccccgtgggctgcacctgcgtccccgagccggagaaggacgcagacagcatcaact 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ccatccccgtgggctgcacctgcgtccccgagccggagaaggacgcagacagcatcaact 420

Qy    421 ccagcatcgacaaacagggcgccaagctcctgctgggccccaacgacgcgcccgctggcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ccagcatcgacaaacagggcgccaagctcctgctgggccccaacgacgcgcccgctggcc 480

Qy    481 cctgaggccggtcctgccccgggaggtctccccggcccgcatcccgaggcgcccaagctg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 cctgaggccggtcctgccccgggaggtctccccggcccgcatcccgaggcgcccaagctg 540

Qy    541 gagccgcctggagggctcggtcggcgacctctgaagagagtgcaccgagcaaaccaagtg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gagccgcctggagggctcggtcggcgacctctgaagagagtgcaccgagcaaaccaagtg 600

Qy    601 ccggagcaccagcgccgcctttccatggagactcgtaagcagcttcatctgacacgggca 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ccggagcaccagcgccgcctttccatggagactcgtaagcagcttcatctgacacgggca 660

Qy    661 tccctggcttgcttttagctacaagcaagcagcgtggctggaagctgatgggaaacgacc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 tccctggcttgcttttagctacaagcaagcagcgtggctggaagctgatgggaaacgacc 720

Qy    721 cggcacgggcatcctgtgtgcggcccgcatggagggtttggaaaagttcacggaggctcc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 cggcacgggcatcctgtgtgcggcccgcatggagggtttggaaaagttcacggaggctcc 780

Qy    781 ctgaggagcctctcagatcggctgctgcgggtgcagggcgtgactcaccgctgggtgctt 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ctgaggagcctctcagatcggctgctgcgggtgcagggcgtgactcaccgctgggtgctt 840

Qy    841 gccaaagagatagggacgcatatgctttttaaagcaatctaaaaataataataagtatag 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gccaaagagatagggacgcatatgctttttaaagcaatctaaaaataataataagtatag 900

Qy    901 cgactatatacctacttttaaaatcaactgttttgaatagaggcagagctattttatatt 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 cgactatatacctacttttaaaatcaactgttttgaatagaggcagagctattttatatt 960

Qy    961 atcaaatgagagctactctgttacatttcttaacatataaacatcgtttttttacttcttc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 atcaaatgagagctactctgttacatttcttaacatataaacatcgtttttttacttcttc 1020

Qy   1021 tggtagaatttttttaaagcataattggaatccttggataaatttttgtagctggtacactc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 tggtagaatttttttaaagcataattggaatccttggataaatttttgtagctggtacactc 1080

Qy   1081 tggcctgggtctctgaattcagcctgtcaccgatggctgactgatgaaatggacacgtct 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 tggcctgggtctctgaattcagcctgtcaccgatggctgactgatgaaatggacacgtct 1140

Qy   1141 catctgacccactcttccttccactgaaggtcttcacgggcctccaggtggaccaaaggg 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 catctgacccactcttccttccactgaaggtcttcacgggcctccaggtggaccaaaggg 1200

Qy   1201 atgcacaggcggctcgcatgccccagggccagctaagagttccaaagatctcagatttgg 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 atgcacaggcggctcgcatgccccagggccagctaagagttccaaagatctcagatttgg 1260

Qy   1261 ttttagtcatgaatacataaacagtctcaaactcgcacaatttttttccccctttgaaag 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ttttagtcatgaatacataaacagtctcaaactcgcacaatttttttccccctttgaaag 1320

Qy   1321 ccactggggccaatttgtggttaagaggtggtgagataagaagtggaacgtgacatcttt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ccactggggccaatttgtggttaagaggtggtgagataagaagtggaacgtgacatcttt 1380

Qy   1381 gccagttgtcagaagaatccaagcaggtattggcttagttgtaagggctttaggatcagg 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 gccagttgtcagaagaatccaagcaggtattggcttagttgtaagggctttaggatcagg 1440

Qy   1441 ctgaatatgaggacaaagtgggccacgttagcatctgcagagatcaatctggaggcttct 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ctgaatatgaggacaaagtgggccacgttagcatctgcagagatcaatctggaggcttct 1500

Qy   1501 gtttctgcattctgccacgagagctaggtccttgatcttttctttagattgaaagtctgt 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 gtttctgcattctgccacgagagctaggtccttgatcttttctttagattgaaagtctgt 1560

Qy   1561 ctctgaacacaattatttgtaaaagttagtagttctttttttaaatcattaaaagaggctt 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ctctgaacacaattatttgtaaaagttagtagttctttttttaaatcattaaaagaggctt 1620
```

```
Qy   1621 gctgaaaaaaaaaaaaaaaaa 1642
          |||||||||||||||||||||
Db   1621 gctgaaaaaaaaaaaaaaaaa 1642


RESULT   2
Z52201
ID   Z52201 standard; cDNA; 2361 BP.
XX
AC   Z52201;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human transforming growth factor beta-9, Ztgfbeta-9 variant cDNA.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             572..1201
FT                   /*tag= a
FT                   /product= "Human Ztgf beta-9 protein"
FT   sig_peptide     76..637
FT                   /*tag= b
FT   mat_peptide     638..1198
FT                   /*tag= c
FT                   /product= "Mature Ztgf beta-9 protein"
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   P-PSDB; Y70662, Y70663.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Example 1; Page 90-92; 97pp; English.
XX
CC   The present sequence encodes a variant of human transforming growth
CC   factor beta-9, designated Ztgf beta-9. Human Ztgf beta-9 was isolated
CC   from an arrayed pituitary gland cDNA plasmid library by PCR screening.
CC   This can be used to treat a variety of neurodegenerative diseases such
CC   as amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's
CC   disease, Parkinson's disease and peripheral neuropathies, or
CC   demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides
CC   have antiviral activity and may also be used to regulate the
CC   proliferation, differentiation and apoptosis of neurons, glial cells,
CC   lymphocytes, hematopoietic cells and stromal cells.
XX
SQ   Sequence 2361 BP; 509 A; 687 C; 655 G; 510 T; 0 other;


  Query Match             98.4%;  Score 1615.6;  DB 21;  Length 2361;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1618;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     21 gccgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 80
          |                                                            |
Db    737 ggcgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 796

Qy     81 gcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    797 gcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 856

Qy    141 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    857 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 916

Qy    201 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    917 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 976

Qy    261 gtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcc 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    977 gtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcc 1036

Qy    321 tgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacc 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1037 tgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacc 1096

Qy    381 tgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggc 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1097 tgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggc 1156

Qy    441 gccaagctcctgctgggccccaacgacgcgcccgctggcccctgaggccggtcctgcccc 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1157 gccaagctcctgctgggccccaacgacgcgcccgctggcccctgaggccggtcctgcccc 1216

Qy    501 gggaggtctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggctcgg 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1217 gggaggtctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggctcgg 1276

Qy    561 tcggcgacctctgaagagagtgcaccgagcaaaccaagtgccggagcaccagcgccgcct 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1277 tcggcgacctctgaagagagtgcaccgagcaaaccaagtgccggagcaccagcgccgcct 1336

Qy    621 ttccatggagactcgtaagcagcttcatctgacacgggcatccctggcttgcttttagct 680
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1337 ttccatggagactcgtaagcagcttcatctgacacgggaatccctggcttgcttttagct 1396

Qy    681 acaagcaagcagcgtggctggaagctgatgggaaacgacccggcacgggcatcctgtgtg 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1397 acaagcaagcagcgtggctggaagctgatgggaaacgacccggcacgggcatcctgtgtg 1456

Qy    741 cggcccgcatggagggtttggaaaagttcacggaggctccctgaggagcctctcagatcg 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1457 cggcccgcatggagggtttggaaaagttcacggaggctccctgaggagcctctcagatcg 1516

Qy    801 gctgctgcgggtgcagggcgtgactcaccgctgggtgcttgccaaagagatagggacgca 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1517 gctgctgcgggtgcagggcgtgactcaccgctgggtgcttgccaaagagatagggacgca 1576

Qy    861 tatgctttttaaagcaatctaaaaataataataagtatagcgactatatacctacttttta 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1577 tatgctttttaaagcaatctaaaaataataataagtatagcgactatatacctacttttta 1636

Qy    921 aaatcaactgttttgaatagaggcagagctattttatattatcaaatgagagctactctg 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1637 aaatcaactgttttgaatagaggcagagctattttatattatcaaatgagagctactctg 1696

Qy    981 ttacatttcttaacatataaacatcgtttttttacttcttctggtagaatttttttaaagca 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1697 ttacatttcttaacatataaacatcgtttttttacttcttctggtagaatttttttaaagca 1756

Qy   1041 taattggaatccttggataaatttttgtagctggtacactctggcctgggtctctgaattc 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1757 taattggaatccttggataaatttttgtagctggtacactctggcctgggtctctgaattc 1816

Qy   1101 agcctgtcaccgatggctgactgatgaaatggacacgtctcatctgacccactcttcctt 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1817 agcctgtcaccgatggctgactgatgaaatggacacgtctcatctgacccactcttcctt 1876

Qy    1161 ccactgaaggtcttcacgggcctccaggtggaccaaagggatgcacaggcggctcgcatg 1220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1877 ccactgaaggtcttcacgggcctccaggtggaccaaagggatgcacaggcggctcgcatg 1936

Qy    1221 ccccagggccagctaagagttccaaagatctcagatttggttttagtcatgaatacataa 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1937 ccccagggccagctaagagttccaaagatctcagatttggttttagtcatgaatacataa 1996

Qy    1281 acagtctcaaactcgcacaatttttttccccctttgaaagccactggggccaatttgtgg 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1997 acagtctcaaactcgcacaatttttttccccctttgaaagccactggggccaatttgtgg 2056

Qy    1341 ttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgtcagaagaatcc 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2057 ttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgtcagaagaatcc 2116

Qy    1401 aagcaggtattggcttagttgtaagggctttaggatcaggctgaatatgaggacaaagtg 1460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2117 aagcaggtattggcttagttgtaagggctttaggatcaggccgaatatgaggacaaagtg 2176

Qy    1461 ggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcattctgccacga 1520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2177 ggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcattctgccacga 2236

Qy    1521 gagctaggtccttgatcttttctttagattgaaagtctgtctctgaacacaattatttgt 1580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2237 gagctaggtccttgatcttttctttagattgaaagtctgtctctgaacacaattatttgt 2296

Qy    1581 aaaagttagtagttctttttttaaatcattaaaagaggcttgctgaaaaaaaaaaaaaaaa 1640
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2297 aaaagttagaagttctttttttaaatcattaaaagaggcttgctgaaaaaaaaaaaaaaaa 2356

Qy    1641 aa 1642
           ||
Db    2357 aa 2358


RESULT   3
Z52195
ID   Z52195 standard; cDNA; 1819 BP.
XX
AC   Z52195;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human transforming growth factor beta-9, Ztgfbeta-9 cDNA.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             71..679
FT                   /*tag=  a
FT                   /product= "Ztgfbeta-9 protein"
FT   sig_peptide     1..115
FT                   /*tag=  b
FT   mat_peptide     116..676
FT                   /*tag=  c
FT                   /product= "Mature Ztgfbeta-9 protein"
FT   mat_peptide     80..676
FT                   /*tag=  d
FT                   /product= "Mature protein represented as SEQ ID NO:3"
FT   mat_peptide     83..676
FT                   /*tag=  e
FT                   /product= "Mature protein represented as SEQ ID NO:4"
FT   mat_peptide     86..676
FT                   /*tag=  f
FT                   /product= "Mature protein represented as SEQ ID NO:5"
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   P-PSDB; Y70653, Y70654, Y70655, Y70656.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 5; Page 81-83; 97pp; English.
XX
CC   The present sequence encodes human transforming growth factor beta-9,
CC   designated Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence 1819 BP; 374 A; 528 C; 530 G; 387 T; 0 other;


  Query Match             97.4%;  Score 1598.6;  DB 21;  Length 1819;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1601;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      21 gccgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 80
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     215 ggcgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 274

Qy      81 gcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     275 gcgagctgcccggcaggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 334

Qy     141 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     335 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 394

Qy     201 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 454

Qy     261 gtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcaccccgcc 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     455 gtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcaccccgcc 514

Qy     321 tgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacc 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     515 tgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacc 574

Qy     381 tgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggc 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     575 tgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggc 634

Qy     441 gccaagctcctgctgggccccaacgacgcgcccgctggcccctgaggccggtcctgcccc 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     635 gccaagctcctgctgggccccaacgacgcgcccgctggcccctgaggccggtcctgcccc 694
```

```
Qy    501 gggaggtctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggctcgg 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    695 gggaggtctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggctcgg 754

Qy    561 tcggcgacctctgaagagagtgcaccgagcaaaccaagtgccggagcaccagcgccgcct 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    755 tcggcgacctctgaagagagtgcaccgagcaaaccaagtgccggagcaccagcgccgcct 814

Qy    621 ttccatggagactcgtaagcagcttcatctgacacgggcatccctggcttgcttttagct 680
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    815 ttccatggagactcgtaagcagcttcatctgacacgggaatccctggcttgcttttagct 874

Qy    681 acaagcaagcagcgtggctggaagctgatgggaaacgacccggcacgggcatcctgtgtg 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    875 acaagcaagcagcgtggctggaagctgatgggaaacgacccggcacgggcatcctgtgtg 934

Qy    741 cggcccgcatggagggtttggaaaagttcacggaggctccctgaggagcctctcagatcg 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    935 cggcccgcatggagggtttggaaaagttcacggaggctccctgaggagcctctcagatcg 994

Qy    801 gctgctgcgggtgcagggcgtgactcaccgctgggtgcttgccaaagagatagggacgca 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    995 gctgctgcgggtgcagggcgtgactcaccgctgggtgcttgccaaagagatagggacgca 1054

Qy    861 tatgctttttaaagcaatctaaaaataataataagtatagcgactatatacctactttta 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1055 tatgctttttaaagcaatctaaaaataataataagtatagcgactatatacctactttta 1114

Qy    921 aaatcaactgttttgaatagaggcagagctattttatattatcaaatgagagctactctg 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1115 aaatcaactgttttgaatagaggcagagctattttatattatcaaatgagagctactctg 1174

Qy    981 ttacatttcttaacatataaacatcgtttttttacttcttctggtagaatttttttaaagca 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1175 ttacatttcttaacatataaacatcgtttttttacttcttctggtagaatttttttaaagca 1234

Qy   1041 taattggaatccttggataaattttgtagctggtacactctggcctgggtctctgaattc 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1235 taattggaatccttggataaattttgtagctggtacactctggcctgggtctctgaattc 1294

Qy   1101 agcctgtcaccgatggctgactgatgaaatggacacgtctcatctgacccactcttcctt 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1295 agcctgtcaccgatggctgactgatgaaatggacacgtctcatctgacccactcttcctt 1354

Qy   1161 ccactgaaggtcttcacgggcctccaggtggaccaaagggatgcacaggcggctcgcatg 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1355 ccactgaaggtcttcacgggcctccaggtggaccaaagggatgcacaggcggctcgcatg 1414

Qy   1221 ccccagggccagctaagagttccaaagatctcagatttggttttagtcatgaatacataa 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1415 ccccagggccagctaagagttccaaagatctcagatttggttttagtcatgaatacataa 1474

Qy   1281 acagtctcaaactcgcacaatttttttccccctttttgaaagccactggggccaatttgtgg 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1475 acagtctcaaactcgcacaatttttttccccctttttgaaagccactggggccaatttgtgg 1534

Qy   1341 ttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgtcagaagaatcc 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1535 ttaagaggtggtgagataagaagtggaacgtgacatctttgccagttgtcagaagaatcc 1594

Qy   1401 aagcaggtattggcttagttgtaagggctttaggatcaggctgaatatgaggacaaagtg 1460
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1595 aagcaggtattggcttagttgtaagggctttaggatcaggccgaatatgaggacaaagtg 1654

Qy   1461 ggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcattctgccacga 1520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1655 ggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcattctgccacga 1714

Qy   1521 gagctaggtccttgatcttttctttagattgaaagtctgtctctgaacacaattatttgt 1580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1715 gagctaggtccttgatcttttctttagattgaaagtctgtctctgaacacaattatttgt 1774

Qy   1581 aaaagttagtagtttctttttttaaatcattaaaagaggcttgctga 1625
          ||||||||| |||||||||||||||||||||||||||||||||||
Db   1775 aaaagttagaagttctttttttaaatcattaaaagaggcttgctga 1819
```

RESULT 4
Z36837
ID Z36837 standard; cDNA; 522 BP.
XX
AC Z36837;
XX
DT 13-MAR-2000 (first entry)
XX
DE Partial nucleotide sequence of human interleukin-22.
XX
KW Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW immune cell chemotaxis; haematopoietic cell disorder;
KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW inflammation; hyperproliferative disorder; tissue regeneration;
KW embryonic stem cell differentiation; embryonic stem cell proliferation;
KW haematopoietic lineage; allergic asthma; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..522
FT /*tag= a
FT /product= "partial IL-22"
XX
PN WO9961617-A1.
XX
PD 02-DEC-1999.
XX
PF 27-MAY-1999; 99WO-US11644.
XX
PR 29-MAY-1998; 98US-0087340.
PR 10-SEP-1998; 98US-0099805.
PR 30-APR-1999; 99US-0131965.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ebner R;
XX
DR WPI; 2000-072622/06.
DR P-PSDB; Y53893.
XX
PT Novel polynucleotides used to develop products for treating e.g. immune
PT disorders, blood disorders, autoimmune disorders, allergies,
PT inflammation, hyperproliferative disorders or infections -
XX
PS Disclosure; Fig 8; 170pp; English.
XX
CC The present sequence encodes a partial human interleukin-22 (IL-22)
CC protein. The specification also describes IL-21 polynucleotides and
CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating
CC deficiencies or disorders of the immune system, by activating or
CC inhibiting the proliferation, differentiation, or mobilization
CC (chemotaxis) of immune cells, treating or detecting deficiencies
CC of haematopoietic cells, to modulate haemostatic or thrombolytic
CC activity, in treating or detecting autoimmune disorders, treating
CC asthma (particularly allergic asthma) or other respiratory problems,
CC to treat and/or prevent organ rejection or graft-versus-host disease
CC (GVHD), to modulate inflammation, to treat or detect hyperproliferative
CC disorders, to treat or detect infectious agents, to differentiate,
CC proliferate and attract cells, leading to the regeneration of tissues,
CC IL-21 and IL-22 may also increase or decrease the differentiation or
CC proliferation of embryonic stem cells and haematopoietic lineage, may
CC be used to modulate mammalian characteristics.
XX
SQ Sequence 522 BP; 77 A; 208 C; 169 G; 68 T; 0 other;

```
  Query Match             28.2%;  Score 463.4;  DB 21;  Length 522;
  Best Local Similarity   99.8%;  Pred. No. 4.4e-93;
  Matches  464;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      21 gccgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 80
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58 ggcgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 117

Qy      81 gcgagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118 gcgagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 177

Qy     141 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     178 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 237

Qy     201 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 260
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     238 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 297

Qy     261 gtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcc 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     298 gtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcc 357

Qy     321 tgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacc 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     358 tgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacc 417

Qy     381 tgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggc 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     418 tgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggc 477

Qy     441 gccaagctcctgctgggccccaacgacgcgcccgctggcccctga 485
           |||||||||||||||||||||||||||||||||||||||||||||
Db     478 gccaagctcctgctgggccccaacgacgcgcccgctggcccctga 522


RESULT   5
Z52198
ID   Z52198 standard; cDNA; 1221 BP.
XX
AC   Z52198;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Murine transforming growth factor beta-9, Ztgfbeta-9 cDNA.
XX
KW   Murine transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             79..696
FT                   /*tag=  a
FT                   /product= "Murine Ztgfbeta-9 protein"
FT   sig_peptide     79..144
FT                   /*tag=  b
FT   mat_peptide     145..693
FT                   /*tag=  c
FT                   /product= "Mature Ztgfbeta-9 protein represented as
FT                   SEQ ID NO:12"
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   P-PSDB; Y70657.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Disclosure; Page 86-88; 97pp; English.
XX
CC   The present sequence encodes murine transforming growth factor beta-9,
CC   designated Ztgf beta-9. Murine Ztgf beta-9 was found to be highly
CC   expressed in the HCL hypothalamic cell line. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence 1221 BP; 224 A; 378 C; 353 G; 266 T; 0 other;


  Query Match             19.9%;  Score 326.4;  DB 21;  Length 1221;
  Best Local Similarity   79.0%;  Pred. No. 7.8e-63;
  Matches  406;  Conservative    0;  Mismatches   96;  Indels   12;  Gaps    1;

Qy      21 gccgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 80
           | |||||||||  ||||||||||||||||||||||  ||||||||  ||||||||||||||
Db     244 ggcgtgctcagcgccttccaccacacgctgcagctcgggccgcgcgagcaggcgcgcaat 303

Qy      81 gcgagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 140
           ||  ||||||||||| |||||||| ||||||||||||||||||||| ||||||||||||
Db     304 gccagctgcccggccgggggcagggccgccgaccgccgcttccggccacccaccaacctg 363

Qy     141 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 200
           |||||||||||||||||||| ||||| || |||||||||| ||  | |  || ||||||
Db     364 cgcagcgtgtcgccctgggcgtacaggatttcctacgaccctgctcgctttccgaggtac 423

Qy     201 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 260
           |||||  |||||||||||||||||| ||||||||||||||||  | ||| |||||||||
Db     424 ctgcccgaagcctactgcctgtgccgaggctgcctgaccgggctctacgggaggaggac 483

Qy     261 gtgcgcttccgcagcgcccctgtctacatgcccaccgtcgtcctgcgccgcacccccgcc 320
           | ||||| |||||| | || |||| |   ||  |||| || ||||| |||||  | |||
Db     484 ttccgctttcgcagcacacccgtcttctctccagccgtggtgctgcggcgcacagcggcc 543

Qy     321 tgcgccggcggccgttccgtctacaccgaggcctacgtcaccatccccgtgggctgcacc 380
           ||||| |||||||| || || ||| ||||   |||| |||||||||| ||||||||||||
Db     544 tgcgcgggcggccgctctgtgtacgccgaacactacatcaccatcccggtgggctgcacc 603

Qy     381 tgcgtccccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggc 440
           ||||| |||||||||||| |||  ||| |||||    |||||||||| ||
Db     604 tgcgtgcccgagccggacaagtccgcggacagtgcgaactccagcatgga---------- 653

Qy     441 gccaagctcctgctgggccccaacgacgcgcccgctggcccctgaggccggtcctgcccc 500
           ||||||  |||||||| |||  ||||  |||  || || | |||| |||||      |||
Db     654 --caagctgctgctggggcccgccgacaggcctgcggggcgctgatgccggggactgccc 711

Qy     501 gggaggtctccccggcccgcatcccgaggcgccc 534
           |  | | | |  |  || |||| |    |  |||
Db     712 gccatggcccagcttcctgcatgcatcaggtccc 745


RESULT   6
Z36853
ID   Z36853 standard; DNA; 498 BP.
```

```
XX
AC   Z36853;
XX
DT   13-MAR-2000  (first entry)
XX
DE   cDNA clone HAGBX04R related to human interleukin-21 CDNA sequence.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma; ss.
XX
OS   Homo sapiens.
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Disclosure; Page 162; 170pp; English.
XX
CC   Z36852-55 represent cDNA clones that are related to cDNA encoding human
CC   interleukin-21 (IL-21) protein. The specification also describes
CC   IL-22 polynucleotides and polypeptides. The IL-21 polynucleotide was
CC   isolated from a cDNA library of apoptotic T-cells. IL-21 and IL-22 may
CC   be useful in treating deficiencies or disorders of the immune system, by
CC   activating or inhibiting the proliferation, differentiation, or
CC   mobilization (chemotaxis) of immune cells, treating or detecting
CC   deficiencies or disorders of haematopoietic cells, to modulate
CC   haemostatic or thrombolytic activity, in treating or detecting
CC   autoimmune disorders, treating asthma (particularly allergic asthma)
CC   or other respiratory problems, to treat and/or prevent organ rejection
CC   or graft-versus-host disease (GVHD), to modulate inflammation, to
CC   treat or detect hyperproliferative disorders, to treat or detect
CC   infectious agents, to differentiate, proliferate and attract cells,
CC   leading to the regeneration of tissues, IL-21 and IL-22 may also
CC   increase or decrease the differentiation or proliferation of embryonic
CC   stem cells and haematopoietic lineage, may be used to modulate mammalian
CC   characteristics.
XX
SQ   Sequence 498 BP; 97 A; 105 C; 147 G; 113 T; 36 other;

  Query Match             14.6%;  Score 239.6;  DB 21;  Length 498;
  Best Local Similarity   75.5%;  Pred. No. 6.9e-44;
  Matches  351;  Conservative    0;  Mismatches  107;  Indels    7;  Gaps    6;

Qy    388 ccgagccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgccaagc 447
          ||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     13 ccagnccggagaaggacgcagacagcatcaactccagcatcgacaaacagggcgccaagc 72

Qy    448 tcctgctgggccccaacgacgcgcccgctggcccctga--ggccggtcctgccccgggag 505
          ||||||||||||||||||||||||||||||||||||  | | | ||||||||||||||
Db     73 tcctgctgggccccaacgacgcgcccgctggcccctgaaggncggttcctgccccgggag 132

Qy    506 gtctccccgg-cccgcatcccgaggcgcccaagctggagccgcctggagggctcggtcgg 564
          |||||||||| ||||||||||||||||||||||||||||||||||||||||  ||||||||
Db    133 gtctccccggncccgcatcccgaggcgcccaagctggagccgcctggaggnttcggtcgg 192

Qy    565 cgacctctgaagagagtgcaccgagcaaaccaagtgccggagcaccagcgccgcc-tttc 623
          ||| |||||||||||||| ||||||||||||||||||||||||| ||||| || | |||
Db    193 cga-ctctgaagagagtncaccgagcaaaccaagtgccggagcaacagcgncgnctttn 251

Qy    624 catggagactcgtaagcag-cttcatctgacacgggcatccctggcttgcttttagctac 682
          |||||||| |||||||||   ||||| ||||| ||| |||||||| ||| |||||| |||
Db    252 catggagattcgtaagcantttttcatttgacangggatccctggtttgtttttagttac 311

Qy    683 aagcaagcagcgtggctggaagctgatgggaaacgacccggcacgggcatcctgtgtgcg 742
          |||||||||   ||| | |||| | |||| || |   | | | ||      | |  | |
Db    312 aagcaagca-nntggnttgaagtngntggggaaaggganccgnagggattctgtnttnggg 370

Qy    743 gcccgcatggagggtttggaaaagttcacggaggctccctgaggagcctctcagatcggc 802
          |||     |  || |||||||||| || | || |   |   | |     |  |
Db    371 gccntntggagggttttggaaaatttnaggggttnctgnggttttttaanattggntt 430

Qy    803 tgctgcgggtgcagggcgtgactcaccgctgggtgcttgccaaag 847
          |  |  ||||  | ||      | ||    | ||  ||   |  |
Db    431 tttttagggttnaagggtnnnttaacttgggngtttttcnaannng 475


RESULT   7
Z36852
ID   Z36852 standard; DNA; 371 BP.
XX
AC   Z36852;
XX
DT   13-MAR-2000  (first entry)
XX
DE   cDNA clone HE2CD08R related to human interleukin-21 CDNA sequence.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma; ss.
XX
OS   Homo sapiens.
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Disclosure; Page 159; 170pp; English.
XX
CC   Z36852-55 represent cDNA clones that are related to cDNA encoding human
CC   interleukin-21 (IL-21) protein. The specification also describes
CC   IL-22 polynucleotides and polypeptides. The IL-21 polynucleotide was
CC   isolated from a cDNA library of apoptotic T-cells. IL-21 and IL-22 may
CC   be useful in treating deficiencies or disorders of the immune system, by
CC   activating or inhibiting the proliferation, differentiation, or
```

```
CC   mobilization (chemotaxis) of immune cells, treating or detecting
CC   deficiencies or disorders of haematopoietic cells, to modulate
CC   haemostatic or thrombolytic activity, in treating or detecting
CC   autoimmune disorders, treating asthma (particularly allergic asthma)
CC   or other respiratory problems, to treat and/or prevent organ rejection
CC   or graft-versus-host disease (GVHD), to modulate inflammation, to
CC   treat or detect hyperproliferative disorders, to treat or detect
CC   infectious agents, to differentiate, proliferate and attract cells,
CC   leading to the regeneration of tissues, IL-21 and IL-22 may also
CC   increase or decrease the differentiation or proliferation of embryonic
CC   stem cells and haematopoietic lineage, may be used to modulate mammalian
CC   characteristics.
XX
SQ   Sequence 371 BP; 98 A; 65 C; 88 G; 106 T; 14 other;


  Query Match             11.9%;  Score 196;  DB 21;  Length 371;
  Best Local Similarity   90.9%;  Pred. No. 2.4e-34;
  Matches  260;  Conservative    0;  Mismatches   21;  Indels    5;  Gaps    5;

Qy    710 gggaaacgacccggcacgggcatcctgtgtgcggcccgcatggagggtttggaaaagttc 769
          | ||||||||||||||| | |   ||||||||||||||||||||||||||||||||||||
Db     15 gngaaacgacccggcangcgatncctgtgtgcggcccgcatggagggtttggaaaagttc 74

Qy    770 acggaggctccctgaggagcctctcagatc-ggctgctgcgggtgcagggcgt-gactca 827
          |||| |||||||||||| | |    ||| |||||||||||||||| ||||| ||||||
Db     75 acggnggctccctgaggacctgcgagaatcgggctgctgcgggtgcaaggcgtggactca 134

Qy    828 ccgctgggtgcttgccaaaga-gatagggacgcatatgctttttaaagcaatctaaaaat 886
          |||||||||||||||||||| | ||||||||||||||||||||||||||||||||||||
Db    135 ccgctgggtgcttgccaaanaggatagggacgcatatgctttttaaagcaatctaaaaat 194

Qy    887 aataataagtatagcgactatatacctacttttaaaatcaactgttttgaatagaggcag 946
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 aataataagtatagcgactatatacctacttttaaaatcaactgttttgaatagaggcag 254

Qy    947 agc-tattttatattat-caaatgagagctactctgttacatttct 990
          ||| |||||||||||| ||||||||||||||||||    || |
Db    255 agcttattttatattatccaaatgagagctactctgttnacatttt 300


RESULT   8
Z36854/c
ID   Z36854 standard; DNA; 178 BP.
XX
AC   Z36854;
XX
DT   13-MAR-2000  (first entry)
XX
DE   cDNA clone HCEBA24FB related to human interleukin-21 CDNA sequence.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma; ss.
XX
OS   Homo sapiens.
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections -
XX
PS   Disclosure; Page 163; 170pp; English.
XX
CC   Z36852-55 represent cDNA clones that are related to cDNA encoding human
CC   interleukin-21 (IL-21) protein. The specification also describes
CC   IL-22 polynucleotides and polypeptides. The IL-21 polynucleotide was
CC   isolated from a cDNA library of apoptotic T-cells. IL-21 and IL-22 may
CC   be useful in treating deficiencies or disorders of the immune system, by
CC   activating or inhibiting the proliferation, differentiation, or
CC   mobilization (chemotaxis) of immune cells, treating or detecting
CC   deficiencies or disorders of haematopoietic cells, to modulate
CC   haemostatic or thrombolytic activity, in treating or detecting
CC   autoimmune disorders, treating asthma (particularly allergic asthma)
CC   or other respiratory problems, to treat and/or prevent organ rejection
CC   or graft-versus-host disease (GVHD), to modulate inflammation, to
CC   treat or detect hyperproliferative disorders, to treat or detect
CC   infectious agents, to differentiate, proliferate and attract cells,
CC   leading to the regeneration of tissues, IL-21 and IL-22 may also
CC   increase or decrease the differentiation or proliferation of embryonic
CC   stem cells and haematopoietic lineage, may be used to modulate mammalian
CC   characteristics.
XX
SQ   Sequence 178 BP; 58 A; 39 C; 31 G; 43 T; 7 other;


  Query Match             10.2%;  Score 167.4;  DB 21;  Length 178;
  Best Local Similarity   95.5%;  Pred. No. 3.6e-28;
  Matches  168;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1451 ggacaaagtgggccacgttagcatctgcagagatcaatctggaggcttctgtttctgcat 1510
           |||||||||||||| |||||||||||| ||| || |||||| ||||||||||||||||
Db    178 GGACAAAGTGGGCCNCGTTAGCATCTGCNGAGNTCNATCTGGNGGCTTCTGTTTCTGCAT 119

Qy   1511 tctgccacgagagctaggtccttgatcttttctttagattgaaagtctgtctctgaacac 1570
           |||||||| ||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    118 TCTGCCACNAGAGCTAGGNCCTTGATCTTTTCTTTAGATTGAAAGTCTGTCTCTGAACAC 59

Qy   1571 aattatttgtaaaagttagtagttctttttaaatcattaaaagaggcttgctgaa 1626
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     58 AATTATTTGTAAAAGTTAGAAGTTCTTTTTTAAATCATTAAAAGAGGCTTGCTGAA 3


RESULT   9
Z36855
ID   Z36855 standard; DNA; 264 BP.
XX
AC   Z36855;
XX
DT   13-MAR-2000  (first entry)
XX
DE   cDNA clone HCELE59R related to human interleukin-21 CDNA sequence.
XX
KW   Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW   immune cell chemotaxis; haematopoietic cell disorder;
KW   haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW   respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW   inflammation; hyperproliferative disorder; tissue regeneration;
KW   embryonic stem cell differentiation; embryonic stem cell proliferation;
KW   haematopoietic lineage; allergic asthma; ss.
XX
OS   Homo sapiens.
XX
PN   WO9961617-A1.
XX
PD   02-DEC-1999.
XX
```

```
PF   27-MAY-1999;   99WO-US11644.
XX
PR   29-MAY-1998;   98US-0087340.
PR   10-SEP-1998;   98US-0099805.
PR   30-APR-1999;   99US-0131965.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ebner R;
XX
DR   WPI; 2000-072622/06.
XX
PT   Novel polynucleotides used to develop products for treating e.g. immune
PT   disorders, blood disorders, autoimmune disorders, allergies,
PT   inflammation, hyperproliferative disorders or infections  -
XX
PS   Disclosure; Page 163-164; 170pp; English.
XX
CC   Z36852-55 represent cDNA clones that are related to cDNA encoding human
CC   interleukin-21 (IL-21) protein. The specification also describes
CC   IL-22 polynucleotides and polypeptides. The IL-21 polynucleotide was
CC   isolated from a cDNA library of apoptotic T-cells. IL-21 and IL-22 may
CC   be useful in treating deficiencies or disorders of the immune system, by
CC   activating or inhibiting the proliferation, differentiation, or
CC   mobilization (chemotaxis) of immune cells, treating or detecting
CC   deficiencies or disorders of haematopoietic cells, to modulate
CC   haemostatic or thrombolytic activity, in treating or detecting
CC   autoimmune disorders, treating asthma (particularly allergic asthma)
CC   or other respiratory problems, to treat and/or prevent organ rejection
CC   or graft-versus-host disease (GVHD), to modulate inflammation, to
CC   treat or detect hyperproliferative disorders, to treat or detect
CC   infectious agents, to differentiate, proliferate and attract cells,
CC   leading to the regeneration of tissues, IL-21 and IL-22 may also
CC   increase or decrease the differentiation or proliferation of embryonic
CC   stem cells and haematopoietic lineage, may be used to modulate mammalian
CC   characteristics.
XX
SQ   Sequence 264 BP; 52 A; 90 C; 77 G; 41 T; 4 other;


  Query Match           10.1%;  Score 165.4;  DB 21;  Length 264;
  Best Local Similarity   87.3%;  Pred. No. 1.1e-27;
  Matches  227;  Conservative    0;  Mismatches   24;  Indels    9;  Gaps    4;

Qy    439 gcgccaagctcctgctgg-gccccaacgacgcgcccgctggcccctgaggccgg-tcctg 496
          | ||||||||||||||||| |||||||||||||||||||||||||||| ||||||| |||||
Db      5 gagccaagctcctgctggngccccaacgacgcgcccgctggcccctaaggccggttcctg 64

Qy    497 ccccgggaggtctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggc 556
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 ccccggaaggtctccccggcccgcatcccgaggcgcccaagctggagccgcctggagggc 124

Qy    557 tcggtcggcgacctctgaagagagtg------caccgagcaaaccaagtgccgg-agcac 609
          | |     |   |    || || |      ||||||||||||||||||||| |||||
Db    125 ttcggtncggcgaacctctgaaagagaagtgccaccgagcaaaccaagtgccggtagcac 184

Qy    610 cagcgccgcctttccatggagactcgtaagcagcttcatctgacacgggcatccctggct 669
          ||| |||||||||||||||||||||||||||||||||||||| ||||| ||||||||| |
Db    185 cagngccgcctttccatggagactcgtaagcagcttcatctganacgggaatccctggtt 244

Qy    670 tgcttttagctacaagcaag 689
          ||||||||||||||||||||
Db    245 tgcttttagctacaagcaag 264


RESULT  10
X89401
ID   X89401 standard; cDNA; 540 BP.
XX
AC   X89401;
XX
DT   25-OCT-1999  (first entry)
XX
DE   Human interleukin-17D polynucleotide.
XX
KW   Interleukin-17D; IL-17D; isoelectric point; cytokine; B cell tumour; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..540
FT                   /*tag=  a
FT                   /product= "IL-17D"
FT                   /note= "Interleukin-17D"
FT   sig_peptide     1..60
FT                   /*tag= b
FT   mat_peptide     61..537
FT                   /*tag= c
XX
PN   WO9935267-A1.
XX
PD   15-JUL-1999.
XX
PF   08-JAN-1999;   99WO-US00513.
XX
PR   09-JAN-1998;   98US-0070886.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Spriggs M,  Upton C;
XX
DR   WPI; 1999-478835/40.
DR   P-PSDB; Y28235.
XX
PT   Murine and Human interleukin 17D DNA, polypeptides and its
PT   fragments, useful as molecular weight markers
XX
PS   Claim 1; Page 6-7; 72pp; English.
XX
CC   This sequence is the human interleukin-17D (IL-17D) polynucleotide.
CC   IL-17D proteins are useful for determining the molecular weight of a
CC   sample protein. Human IL-17D polypeptides are significantly related to
CC   human IL-17 polypeptides. The homology between IL-17 and IL-17D suggests
CC   that the IL-17D polypeptide is capable of signalling through cytokine
CC   receptors. The IL-17D protein and fragments of it are useful as controls
CC   for peptide fragmentation which can be used to determine the isoelectric
CC   point of a sample protein. Antibodies generated against IL-17D and its
CC   fragmented peptides can be used to enhance the accuracy of these
CC   molecular weight markers. IL-17D can also be used as a therapeutic agent
CC   for the treatment of diseases mediated by IL-17D. IL-17D polypeptides
CC   bind to B cells. It is likely that these polypeptides can be used for
CC   targeting compounds to B cells and B cell tumours, and for specific
CC   selection of B cell populations.
XX
SQ   Sequence 540 BP; 102 A; 174 C; 169 G; 95 T; 0 other;


  Query Match           4.5%;  Score 73.2;  DB 20;  Length 540;
  Best Local Similarity   58.1%;  Pred. No. 2.7e-07;
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy    133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
          |||||  | | ||| |||| ||||||| |||||| |||  | |||||||     |   ||
Db    278 ccaacaagaggagcctgtctccctggggctacagcatcaaccacgaccccagccgtatcc 337

Qy    193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
          ||  | |||||||| || ||    |||||||| | ||||||  ||| |    |   |
Db    338 ccgtggacctgccggaggcacggtgcctgtgtctgggctgtgtgaacccccttcaccatgc 397

Qy    253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
          ||||||||    || |    ||||  || || | ||        | ||       |  ||  |
Db    398 aggaggaccgcagcatggtgagcgtgccggtgttcagccaggttcctgtgcgccgccgcc 457

Qy    307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
            || | ||  |   |  |   || | || |  | |   ||  | |   |  ||||||
Db    458 tctgcccgccaccgccccgcacagggccttgccgccagcgcgcagtcatggagaccatcg 517
```

```
Qy    367 ccgtgggctgcacctgcgtc 386
          | ||||||||||||||| ||
Db    518 ctgtgggctgcacctgcatc 537


RESULT 11
Z34282
ID   Z34282 standard; cDNA; 687 BP.
XX
AC   Z34282;
XX
DT   07-DEC-1999  (first entry)
XX
DE   Human PRO1031 nucleotide sequence.
XX
KW   Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW   probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW   secreted protein; transmembrane protein; ss.
XX
OS   Homo sapiens.
XX
PN   WO9946281-A2.
XX
PD   16-SEP-1999.
XX
PF   08-MAR-1999;   99WO-US05028.
XX
PR   10-MAR-1998;   98US-0077450.
PR   11-MAR-1998;   98US-0077632.
PR   11-MAR-1998;   98US-0077641.
PR   11-MAR-1998;   98US-0077649.
PR   12-MAR-1998;   98US-0077791.
PR   13-MAR-1998;   98US-0078004.
PR   17-MAR-1998;   98US-0040220.
PR   20-MAR-1998;   98US-0078886.
PR   20-MAR-1998;   98US-0078910.
PR   20-MAR-1998;   98US-0078936.
PR   20-MAR-1998;   98US-0078939.
PR   25-MAR-1998;   98US-0079294.
PR   26-MAR-1998;   98US-0079656.
PR   27-MAR-1998;   98US-0079663.
PR   27-MAR-1998;   98US-0079664.
PR   27-MAR-1998;   98US-0079689.
PR   27-MAR-1998;   98US-0079728.
PR   27-MAR-1998;   98US-0079786.
PR   30-MAR-1998;   98US-0079920.
PR   30-MAR-1998;   98US-0079923.
PR   31-MAR-1998;   98US-0080105.
PR   31-MAR-1998;   98US-0080107.
PR   31-MAR-1998;   98US-0080165.
PR   31-MAR-1998;   98US-0080194.
PR   01-APR-1998;   98US-0080327.
PR   01-APR-1998;   98US-0080328.
PR   01-APR-1998;   98US-0080333.
PR   01-APR-1998;   98US-0080334.
PR   08-APR-1998;   98US-0081049.
PR   08-APR-1998;   98US-0081070.
PR   08-APR-1998;   98US-0081071.
PR   09-APR-1998;   98US-0081195.
PR   09-APR-1998;   98US-0081203.
PR   09-APR-1998;   98US-0081229.
PR   15-APR-1998;   98US-0081817.
PR   15-APR-1998;   98US-0081838.
PR   15-APR-1998;   98US-0081952.
PR   15-APR-1998;   98US-0081955.
PR   21-APR-1998;   98US-0082568.
PR   21-APR-1998;   98US-0082569.
PR   22-APR-1998;   98US-0082700.
PR   22-APR-1998;   98US-0082704.
PR   22-APR-1998;   98US-0082804.
PR   23-APR-1998;   98US-0082767.
PR   23-APR-1998;   98US-0082796.
PR   27-APR-1998;   98US-0083336.
PR   28-APR-1998;   98US-0083322.
PR   29-APR-1998;   98US-0083392.
PR   29-APR-1998;   98US-0083495.
PR   29-APR-1998;   98US-0083496.
PR   29-APR-1998;   98US-0083499.
PR   29-APR-1998;   98US-0083500.
PR   29-APR-1998;   98US-0083545.
PR   29-APR-1998;   98US-0083554.
PR   29-APR-1998;   98US-0083558.
PR   29-APR-1998;   98US-0083559.
PR   30-APR-1998;   98US-0083742.
PR   05-MAY-1998;   98US-0084366.
PR   06-MAY-1998;   98US-0084414.
PR   06-MAY-1998;   98US-0084441.
PR   07-MAY-1998;   98US-0084598.
PR   07-MAY-1998;   98US-0084600.
PR   07-MAY-1998;   98US-0084627.
PR   07-MAY-1998;   98US-0084637.
PR   07-MAY-1998;   98US-0084639.
PR   07-MAY-1998;   98US-0084640.
PR   07-MAY-1998;   98US-0084643.
PR   13-MAY-1998;   98US-0085323.
PR   13-MAY-1998;   98US-0085338.
PR   13-MAY-1998;   98US-0085339.
PR   15-MAY-1998;   98US-0085573.
PR   15-MAY-1998;   98US-0085579.
PR   15-MAY-1998;   98US-0085580.
PR   15-MAY-1998;   98US-0085582.
PR   15-MAY-1998;   98US-0085689.
PR   15-MAY-1998;   98US-0085697.
PR   15-MAY-1998;   98US-0085700.
PR   15-MAY-1998;   98US-0085704.
PR   18-MAY-1998;   98US-0086023.
PR   22-MAY-1998;   98US-0086392.
PR   22-MAY-1998;   98US-0086414.
PR   22-MAY-1998;   98US-0086430.
PR   22-MAY-1998;   98US-0086486.
PR   28-MAY-1998;   98US-0087098.
PR   28-MAY-1998;   98US-0087106.
PR   28-MAY-1998;   98US-0087208.
PR   30-JUL-1998;   98US-0094651.
PR   11-SEP-1998;   98US-0100038.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wood WI,  Goddard A,  Gurney A,  Yuan J,  Baker KP,  Chen J;
XX
DR   WPI; 1999-551358/46.
DR   P-PSDB; Y41762.
XX
PT   New secreted and transmembrane polypeptides and their polynucleotides,
PT   useful for treating blood coagulation disorders, cancers and cellular
PT   adhesion disorders -
XX
PS   Claim 2; Fig 197; 530pp; English.
XX
CC   The present invention describes secreted and transmembrane polypeptides
CC   and their polynucleotides. The nucleotide sequences are useful as
CC   sources of probes, primers, for chromosome mapping, and for generation
CC   of antisense sequences. They can also be used to create transgenic
CC   animals. The proteins can be used to treat a variety of diseases and
CC   disorders, depending on their function. Diseases that may be treated
CC   include blood coagulation disorders, cancers and cellular adhesion
CC   disorders. They may also be used to raise antibodies. Z33891 to
CC   Z34338, and Y41685 to Y41774 represent polynucleotide and polypeptide
CC   sequence given in the exemplification of the present invention.
XX
SQ   Sequence 687 BP; 142 A; 216 C; 209 G; 120 T; 0 other;


  Query Match             4.5%;  Score 73.2;  DB 20;  Length 687;
  Best Local Similarity   58.1%;  Pred. No. 3e-07;
```

```
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy    133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
          |||||  | | ||| |||| |||||||| |||||| |||  | |||||||     |   ||
Db    319 ccaacaagaggagcctgtctccctggggctacagcatcaaccacgaccccagccgtatcc 378

Qy    193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
          ||  | ||||||| || ||    |||||||| | ||||||  ||| |     |   |
Db    379 ccgtggacctgccggaggcacggtgcctgtgtctgggctgtgtgaaccccttcaccatgc 438

Qy    253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
          ||||||||     || |    ||||  || || | ||      | ||      |  ||  |
Db    439 aggaggaccgcagcatggtgagcgtgccggtgttcagccaggttcctgtgcgccgccgcc 498

Qy    307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
           | || | ||  |   |  |   || | || |  | |   ||  | |   |  ||||||
Db    499 tctgcccgccaccgccccgcacagggccttgccgccagcgcgcagtcatggagaccatcg 558

Qy    367 ccgtgggctgcacctgcgtc 386
          | ||||||||||||||| ||
Db    559 ctgtgggctgcacctgcatc 578


RESULT  12
Z29727
ID   Z29727 standard; cDNA; 687 BP.
XX
AC   Z29727;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Human Interleukin 17B, PRO1031 cDNA.
XX
KW   Interleukin; IL-17B; PRO1031 polypeptide; clone DNA59294-1381; UNQ516;
KW   cytokine IL-17; cytotoxic T-lymphocyte-associated antigen 8; CTLA-8;
KW   hybridisation probe; antagonist; degenerative cartilaginous disorder;
KW   agonist; diagnose; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             42..584
FT                   /*tag=  a
FT                   /product= "Interleukin 17B-PRO1031 polypeptide"
FT                   /note= "Designated as clone DNA59294-1381"
FT   sig_peptide     42..101
FT                   /*tag=  b
FT   mat_peptide     102..581
FT                   /*tag=  c
FT                   /label= Mature_Interleukin_17B_polypeptide
FT                   /note= "Used to treat degenerative cartilaginous
FT                   disorder"
XX
PN   WO9960127-A2.
XX
PD   25-NOV-1999.
XX
PF   14-MAY-1999;   99WO-US10733.
XX
PR   15-MAY-1998;   98US-0085579.
PR   23-DEC-1998;   98US-0113621.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Filvaroff E,  Goddard A,  Gurney AL,  Li H,  Wood WI;
XX
DR   WPI; 2000-116314/10.
DR   P-PSDB; Y44459.
XX
PT   New polypeptides designated PRO1031 and PRO1122 used to treat a
PT   degenerative cartilaginous disorder -
XX
PS   Claim 2; Fig 2; 141pp; English.
XX
CC   The present sequence is the cDNA clone DNA59294-1381, encoding the
CC   human PRO1031 polypeptide, also referred to as UNQ516, and
CC   as interleukin-17B (IL-17B). This sequence has identity with the
CC   cytokine IL-17 and cytotoxic T-lymphocyte-associated antigen 8 (CTLA-8).
CC   PRO1031 is expressed in pancreas, small intestine, stomach and testis
CC   also. It shares about 26-28% amino acid identity with IL-17 and IL-17C.
CC   The entire coding region of IL-17B can be used as hybridisation probe.
CC   The PRO1031 polypeptide, agonist or antagonist, is used to diagnose and
CC   treat a degenerative cartilaginous disorder.
XX
SQ   Sequence 687 BP; 142 A; 216 C; 209 G; 120 T; 0 other;


  Query Match             4.5%;  Score 73.2;  DB 21;  Length 687;
  Best Local Similarity   58.1%;  Pred. No. 3e-07;
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy    133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
          |||||  | | ||| |||| |||||||| |||||| |||  | |||||||     |   ||
Db    319 ccaacaagaggagcctgtctccctggggctacagcatcaaccacgaccccagccgtatcc 378

Qy    193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
          ||  | ||||||| || ||    |||||||| | ||||||  ||| |     |   |
Db    379 ccgtggacctgccggaggcacggtgcctgtgtctgggctgtgtgaaccccttcaccatgc 438

Qy    253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
          ||||||||     || |    ||||  || || | ||      | ||      |  ||  |
Db    439 aggaggaccgcagcatggtgagcgtgccggtgttcagccaggttcctgtgcgccgccgcc 498

Qy    307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
           | || | ||  |   |  |   || | || |  | |   ||  | |   |  ||||||
Db    499 tctgcccgccaccgccccgcacagggccttgccgccagcgcgcagtcatggagaccatcg 558

Qy    367 ccgtgggctgcacctgcgtc 386
          | ||||||||||||||| ||
Db    559 ctgtgggctgcacctgcatc 578


RESULT  13
X15870
ID   X15870 standard; cDNA; 705 BP.
XX
AC   X15870;
XX
DT   11-MAY-1999  (first entry)
XX
DE   cDNA encoding interleukin-20.
XX
KW   Human interleukin-20; IL-20; B-cell neoplasm; chronic lymphocyte
KW   leukemia; B-lymphocyte leukemia; anticancer; antiviral;
KW   immunodeficiency; rheumatoid arthritis; antiinflammation; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             45..587
FT                   /*tag=  a
FT                   /product= IL-20
XX
PN   WO9903982-A1.
XX
PD   28-JAN-1999.
XX
PF   15-JUL-1998;   98WO-US14609.
XX
PR   26-SEP-1997;   97US-0060140.
PR   16-JUL-1997;   97US-0052870.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Duan RD,  Ebner R,  Florence KA,  Hu J,  Ruben SM;
XX
```

```
DR   WPI; 1999-132229/11.
DR   P-PSDB; W97350.
XX
PT   New interleukin-20 (IL-20) protein  - used to detect and treat
PT   immune system related disorders
XX
PS   Claim 2; Fig 1; 80pp; English.
XX
CC   The present sequence encodes human interleukin-20 (IL-20). Il-20
CC   may be used to treat B-cell neoplasms, including chronic lymphocyte
CC   leukemia (CLL) and B-lymphocyte leukemia (BLL), and in anticancer and
CC   antiviral treatments. Immunodeficiencies, e.g. in T- and B-lymphocytes,
CC   leukopenia, reduced numbers of leukocytes, immune disorders,
CC   e.g. rhematoid arthritis, may be treated with IL-20. IL-20 may also be
CC   used to augment the humoral or cellular immune response in vivo
CC   in coadministration with other therapeutic agents, e.g. to enhance the
CC   efficacy of viral antigen vaccines, such as HIV. IL-20 may also be
CC   useful in immunotherapeutic and antiinflammation compositions, for the
CC   treatment of patients suffering from chemotherapy from bone marrow
CC   transplants, to treat corneal damage, keratitis, ulcers,
CC   thrombocytopenia, to restore neutrophil and platelet counts in treatment
CC   of cancer, to enhance erythropoietic production for treating anaemias
CC   associated with inflammation, renal failure, AIDS and cancer. IL-20 may
CC   be employed to treat hematopoiesis, and to treat sepsis. Agonists and
CC   antagonists of IL-20 can also be used.
XX
SQ   Sequence 705 BP; 158 A; 217 C; 208 G; 122 T; 0 other;


  Query Match             4.5%;  Score 73.2;  DB 20;  Length 705;
  Best Local Similarity   58.1%;  Pred. No. 3e-07;
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy    133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
          |||||  | | ||| |||| ||||||| |||||| |||  | |||||||     |   ||
Db    322 ccaacaagaggagcctgtctccctggggctacagcatcaaccacgaccccagccgtatcc 381

Qy    193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
          ||  | |||||||  || ||    |||||||| | ||||||  ||| |    |   |
Db    382 ccgtggacctgccggaggcacggtgcctgtgtctgggctgtgtgaacccccttcaccatgc 441

Qy    253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
          ||||||||    || |    ||||  || || | ||      | ||      |  ||  |
Db    442 aggaggaccgcagcatggtgagcgtgccggtgttcagccaggttcctgtgcgccgccgcc 501

Qy    307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
           | || | ||  |   |  |   || | || |  | |   ||  | |   |  ||||||
Db    502 tctgcccgccaccgccccgcacagggccttgccgccagcgcgcagtcatggagaccatcg 561

Qy    367 ccgtgggctgcacctgcgtc 386
          | ||||||||||||||| ||
Db    562 ctgtgggctgcacctgcatc 581


RESULT  14
V83172
ID   V83172 standard; cDNA; 736 BP.
XX
AC   V83172;
XX
DT   02-MAR-1999  (first entry)
XX
DE   Human Zcyto7 cDNA.
XX
KW   Proliferation; differentiation; hormone; growth factor; interleukin;
KW   cytokine; cytokine-like factor; spinal cord; glial cell; neuron;
KW   neurodegenerative disease; multiple sclerosis; autoimmune disease;
KW   cancer; leukaemia; inflammation; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             57..599
FT                   /*tag=  a
FT                   /product= Zcyto7
XX
PN   WO9849310-A1.
XX
PD   05-NOV-1998.
XX
PF   23-APR-1998;   98WO-US08212.
XX
PR   16-JAN-1998;   98US-0071676.
PR   25-APR-1997;   97US-0044886.
PR   25-APR-1997;   97US-0845753.
PR   16-JAN-1998;   98US-0008543.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Foster DC,  Gilbert T,  Hart CE,  Hoffmann RC,  Lehner JM;
PI   Martinez T,  Presnell SR,  Whitmore TE;
XX
DR   WPI; 1999-024061/02.
DR   P-PSDB; W85615.
XX
PT   Mammalian cytokine-like factor 7 polypeptide, Zcyto7 - useful e.g.
PT   to promote bone and cartilage growth e.g. to treat osteoporosis, or
PT   in treatment of inflammation, neurodegenerative diseases etc.
XX
PS   Claim 1; Page 53-54; 87pp; English.
XX
CC   Zcyto7 is useful therapeutically to promote bone and cartilage growth
CC   e.g. to treat osteoporosis or osteoarthritis or stimulate
CC   regeneration of dentin/bone lost due to periodontal disease. Zcyto7
CC   may be involved in spinal cord maintenance and may be useful to treat
CC   neurodegenerative diseases (e.g. amyotrophic lateral sclerosis),
CC   multiple sclerosis or sensory neuropathies.  It may also be useful to
CC   treat inflammation or modulate activity of immune system cells (e.g.
CC   to treat autoimmune diseases), or to help regulate other cytokines in
CC   the spinal cord.  It can be used to identify antagonists or agonists
CC   and to produce antibodies useful therapeutically or diagnostically.
CC   The antibodies are also useful to purify Zcyto7 and manufacture
CC   anti-idiotypic antibodies.  Polynucleotides encoding Zcyto7 can be
CC   introduced therapeutically into mammalian cells lacking the Zcyto7
CC   gene/having a mutated gene or used diagnostically to detect gene
CC   mutations.
XX
SQ   Sequence 736 BP; 166 A; 226 C; 219 G; 125 T; 0 other;


  Query Match             4.5%;  Score 73.2;  DB 20;  Length 736;
  Best Local Similarity   58.1%;  Pred. No. 3e-07;
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy    133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
          |||||  | | ||| |||| ||||||| |||||| |||  | |||||||     |   ||
Db    334 ccaacaagaggagcctgtctccctggggctacagcatcaaccacgaccccagccgtatcc 393

Qy    193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
          ||  | |||||||  || ||    |||||||| | ||||||  ||| |    |   |
Db    394 ccgtggacctgccggaggcacggtgcctgtgtctgggctgtgtgaacccccttcaccatgc 453

Qy    253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
          ||||||||    || |    ||||  || || | ||      | ||      |  ||  |
Db    454 aggaggaccgcagcatggtgagcgtgccggtgttcagccaggttcctgtgcgccgccgcc 513

Qy    307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
           | || | ||  |   |  |   || | || |  | |   ||  | |   |  ||||||
Db    514 tctgcccgccaccgccccgcacagggccttgccgccagcgcgcagtcatggagaccatcg 573

Qy    367 ccgtgggctgcacctgcgtc 386
          | ||||||||||||||| ||
Db    574 ctgtgggctgcacctgcatc 593


RESULT  15
```

```
X84472
ID   X84472 standard; DNA; 738 BP.
XX
AC   X84472;
XX
DT   10-SEP-1999  (first entry)
XX
DE   Human EDIRF I coding sequence.
XX
KW   Embryo derived interleukin related factor; diagnosis; detection; therapy;
KW   EDIRF-related disease; immune disorder; haematopoietic disorder;
KW   developmental disorder; inflammatory disease; arthritis; psoriasis;
KW   EDIRF I; TANGO89; ss.
XX
OS   Homo sapiens.
XX
PN   WO9932632-A1.
XX
PD   01-JUL-1999.
XX
PF   18-DEC-1998;   98WO-US27068.
XX
PR   19-DEC-1997;   97US-0994890.
XX
PA   (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI   Holtzman DA;
XX
DR   WPI; 1999-418929/35.
DR   P-PSDB; Y22197.
XX
PT   Nucleic acid encoding embryo-derived interleukin-related factors
XX
PS   Claim 1; Fig 1a; 116pp; English.
XX
CC   This sequence encodes a embryo-derived interleukin-related factor (EDIRF)
CC   of the invention, designated human EDIRF I (also known as TANGO89).
CC   The EDIRF DNA and protein sequences (and their homologues), antibodies
CC   (Ab) specific for EDIRF, and other modulators are used: (i) in screening
CC   and detection assays, e.g. for chromosome mapping, tissue typing or
CC   forensic studies; (ii) in diagnosis, prognosis or monitoring clinical
CC   trials; and (iii) for treating or preventing EDIRF-related diseases
CC   (especially immune, haematopoietic, differentiative, developmental or
CC   inflammatory disease, including arthritis and psoriasis. The EDIRF coding
CC   sequence, or its fragments, are also useful as probes and primers (for
CC   detecting related sequences and disease-associated mutations, also for
CC   mutagenesis), for expressing recombinant EDIRF and as source of
CC   antisense, ribozyme and peptide nucleic acids for inhibiting translation
CC   of EDIRF-derived mRNA. EDIRF is used to raise Ab (useful for detecting
CC   EDIRF, including forms with aberrant post-translational modification, for
CC   affinity purification and therapeutically) and to screen for specific
CC   modulators (e.g. peptides or peptidomimetics).
XX
SQ   Sequence 738 BP; 161 A; 230 C; 222 G; 125 T; 0 other;


  Query Match             4.5%;  Score 73.2;  DB 20;  Length 738;
  Best Local Similarity   58.1%;  Pred. No. 3e-07;
  Matches  151;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    1;

Qy     133 ccaacctgcgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtacc 192
           |||||  | | ||| |||| |||||||| |||||| |||  | |||||||     |   ||
Db     346 ccaacaagaggagcctgtctccctggggctacagcatcaaccacgaccccagccgtatcc 405

Qy     193 ccaggtacctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcg 252
           ||  | |||||||| || ||    |||||||| | ||||||  ||| |     |   |
Db     406 ccgtggacctgccggaggcacggtgcctgtgtctgggctgtgtgaacccccttcaccatgc 465

Qy     253 aggaggacgtgcgcttccgcagcgcccctgtctaca------tgcccaccgtcgtcctgc 306
           ||||||||    || |    ||||  || || | ||      | ||      |  ||  |
Db     466 aggaggaccgcagcatggtgagcgtgccggtgttcagccaggttcctgtgcgccgccgcc 525

Qy     307 gccgcacccccgcctgcgccggcggccgttccgtctacaccgaggcctacgtcaccatcc 366
           | || | ||  |   |  |   || | || |  | |   ||  | |   |  ||||||
Db     526 tctgcccgccaccgccccgcacagggccttgccgccagcgcgcagtcatggagaccatcg 585

Qy     367 ccgtgggctgcacctgcgtc 386
           | ||||||||||||||| ||
Db     586 ctgtgggctgcacctgcatc 605


Search completed: February 24, 2001, 21:51:44
Job time: 7096 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:49:48 ; Search time 26.41 Seconds
(without alignments)
207.157 Million cell updates/sec

Title: US-09-320-713-4
Perfect score: 864
Sequence: 1 NSARARAVLSAFHHTLQLGP.....SIDKQGAKLLLGPNDAPAGP 160

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	864	100.0	160	21	Y53891	Partial amino acid
2	835	96.6	173	21	Y53893	Partial amino acid
3	835	96.6	185	21	Y70656	Mature human trans
4	835	96.6	186	21	Y70655	Mature human trans
5	835	96.6	187	21	Y70654	Mature human trans
6	835	96.6	187	21	Y70663	Mature human Ztgfb
7	835	96.6	202	21	Y70653	Human transforming
8	835	96.6	209	21	Y70662	Human transforming
9	707	81.8	183	21	Y70658	Mature murine tran
10	707	81.8	205	21	Y70657	Murine transformin
11	303	35.1	57	21	Y70666	Human transforming
12	298	34.5	54	21	Y70664	Human transforming
13	197.5	22.9	425	21	Y44462	Human Interleukin
14	192	22.2	197	21	Y92238	Human interleukin-
15	192	22.2	197	21	Y44460	Human Interleukin
16	192	22.2	197	21	Y53892	Amino acid sequenc
17	192	22.2	206	21	Y44485	Human Interleukin
18	180	20.8	34	21	Y70661	Human transforming
19	177	20.5	158	20	W85646	Murine Zcyto7 matu
20	177	20.5	160	20	W85645	Murine Zcyto7 matu
21	177	20.5	180	20	Y28236	Mouse interleukin-
22	177	20.5	180	20	Y22198	Murine EDIRF I pro
23	177	20.5	180	20	W85617	Murine Zcyto7. Mu
24	176.5	20.4	408	21	Y44461	Human Interleukin
25	174	20.1	128	20	W85643	Human Zcyto7 matur
26	174	20.1	151	20	W85641	Human Zcyto7 matur
27	174	20.1	153	20	W85642	Human Zcyto7 matur
28	174	20.1	154	20	W85640	Human Zcyto7 matur
29	174	20.1	157	20	W85644	Human Zcyto7 matur
30	174	20.1	158	20	W85639	Human Zcyto7 matur
31	174	20.1	160	20	W85618	Human Zcyto7 matur
32	174	20.1	160	20	W85619	Human Zcyto7 matur
33	174	20.1	160	20	W85620	Human Zcyto7 matur
34	174	20.1	160	20	W85621	Human Zcyto7 matur
35	174	20.1	160	20	W85622	Human Zcyto7 matur
36	174	20.1	160	20	W85626	Human Zcyto7 matur
37	174	20.1	160	20	W85627	Human Zcyto7 matur
38	174	20.1	160	20	W85628	Human Zcyto7 matur
39	174	20.1	160	20	W85629	Human Zcyto7 matur
40	174	20.1	175	21	Y44484	Human Interleukin
41	174	20.1	180	20	Y41762	Human PRO1031 prot
42	174	20.1	180	20	Y28235	Human interleukin-
43	174	20.1	180	20	Y22197	Human EDIRF I prot
44	174	20.1	180	20	W97350	Interleukin-20. H
45	174	20.1	180	20	W85615	Human Zcyto7. Hom

ALIGNMENTS

RESULT 1
Y53891
ID Y53891 standard; Protein; 160 AA.
XX
AC Y53891;
XX
DT 13-MAR-2000 (first entry)
XX
DE Partial amino acid sequence of human interleukin-22.
XX
KW Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW immune cell chemotaxis; haematopoietic cell disorder;
KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW inflammation; hyperproliferative disorder; tissue regeneration;
KW embryonic stem cell differentiation; embryonic stem cell proliferation;
KW haematopoietic lineage; allergic asthma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 26
FT /note= "potential N-linked glycosylation site"
FT Domain 57..64
FT /note= "conserved domain I"
FT Domain 72..77
FT /note= "conserved domain II"
FT Domain 99..105
FT /note= "conserved domain III"
FT Domain 121..128
FT /note= "conserved domain IV"
FT Modified-site 139
FT /note= "potential N-linked glycosylation site"
XX
PN WO9961617-A1.

XX
PD 02-DEC-1999.
XX
PF 27-MAY-1999; 99WO-US11644.
XX
PR 29-MAY-1998; 98US-0087340.
PR 10-SEP-1998; 98US-0099805.
PR 30-APR-1999; 99US-0131965.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ebner R;
XX
DR WPI; 2000-072622/06.
DR N-PSDB; Z36835.
XX
PT Novel polynucleotides used to develop products for treating e.g. immune
PT disorders, blood disorders, autoimmune disorders, allergies,
PT inflammation, hyperproliferative disorders or infections -
XX
PS Claim 27; Fig 2A-B; 170pp; English.
XX
CC The present sequence represents a partial human interleukin-22 (IL-22)
CC protein. The specification also describes IL-21 polynucleotides and
CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating
CC deficiencies or disorders of the immune system, by activating or
CC inhibiting the proliferation, differentiation, or mobilization
CC (chemotaxis) of immune cells, treating or detecting deficiencies or
CC disorders of haematopoietic cells, to modulate haemostatic or
CC thrombolytic activity, in treating or detecting autoimmune disorders,
CC treating asthma (particularly allergic asthma) or other respiratory
CC problems, to treat and/or prevent organ rejection or graft-versus-host
CC disease (GVHD), to modulate inflammation, to treat or detect
CC hyperproliferative disorders, to treat or detect infectious agents, to
CC differentiate, proliferate and attract cells, leading to the
CC regeneration of tissues, IL-21 and IL-22 may also increase or decrease
CC the differentiation or proliferation of embryonic stem cells and
CC haematopoietic lineage, may be used to modulate mammalian
CC characteristics.
XX
SQ Sequence 160 AA;

Query Match 100.0%; Score 864; DB 21; Length 160;
Best Local Similarity 100.0%; Pred. No. 4.3e-89;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 NSARARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 nsararavlsafhhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydp 60

Qy      61 ARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aryprylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsvyteayvt 120

Qy     121 IPVGCTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAGP 160
           ||||||||||||||||||||||||||||||||||||||||
Db     121 ipvgctcvpepekdadsinssidkqgakllgpndapagp 160
```

RESULT 2
Y53893
ID Y53893 standard; Protein; 173 AA.
XX
AC Y53893;
XX
DT 13-MAR-2000 (first entry)
XX
DE Partial amino acid sequence of human interleukin-22.
XX
KW Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW immune cell chemotaxis; haematopoietic cell disorder;
KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW inflammation; hyperproliferative disorder; tissue regeneration;
KW embryonic stem cell differentiation; embryonic stem cell proliferation;
KW haematopoietic lineage; allergic asthma.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Domain	18..23
FT		/note= "conserved domain VI"
FT	Modified-site	39..41
FT		/note= "Asn39 is a potential N-linked glycosylation site"
FT	Domain	60..65
FT		/note= "conserved domain VII"
FT	Domain	69..77
FT		/note= "conserved domain I"
FT	Domain	85..90
FT		/note= "conserved domain II"
FT	Domain	112..118
FT		/note= "conserved domain III"
FT	Domain	134..141
FT		/note= "conserved domain IV"
FT	Modified-site	152..154
FT		/note= "Asn152 is a potential N-linked glycosylation site"

XX
PN WO9961617-A1.
XX
PD 02-DEC-1999.
XX
PF 27-MAY-1999; 99WO-US11644.
XX
PR 29-MAY-1998; 98US-0087340.
PR 10-SEP-1998; 98US-0099805.
PR 30-APR-1999; 99US-0131965.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ebner R;
XX
DR WPI; 2000-072622/06.
DR N-PSDB; Z36837.
XX
PT Novel polynucleotides used to develop products for treating e.g. immune
PT disorders, blood disorders, autoimmune disorders, allergies,
PT inflammation, hyperproliferative disorders or infections -
XX
PS Disclosure; Fig 8; 170pp; English.
XX
CC The present sequence represents a partial human interleukin-22 (IL-22)
CC protein. The specification also describes IL-21 polynucleotides and
CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library
CC of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating
CC deficiencies or disorders of the immune system, by activating or
CC inhibiting the proliferation, differentiation, or mobilization
CC (chemotaxis) of immune cells, treating or detecting deficiencies
CC of haematopoietic cells, to modulate haemostatic or thrombolytic
CC activity, in treating or detecting autoimmune disorders, treating
CC asthma (particularly allergic asthma) or other respiratory problems,
CC to treat and/or prevent organ rejection or graft-versus-host disease
CC (GVHD), to modulate inflammation, to treat or detect hyperproliferative
CC disorders, to treat or detect infectious agents, to differentiate,
CC proliferate and attract cells, leading to the regeneration of tissues,
CC IL-21 and IL-22 may also increase or decrease the differentiation or
CC proliferation of embryonic stem cells and haematopoietic lineage, may
CC be used to modulate mammalian characteristics.
XX
SQ Sequence 173 AA;

Query Match 96.6%; Score 835; DB 21; Length 173;
Best Local Similarity 98.7%; Pred. No. 8.2e-86;

```
  Matches  154;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18 aagvlsafhhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyp 77

Qy      65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78 rylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsvyteayvtipvg 137

Qy     125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAGP 160
           ||||||||||||||||||||||||||||||||||||
Db     138 ctcvpepekdadsinssidkqgakllgpndapagp 173


RESULT   3
Y70656
ID   Y70656 standard; Protein; 185 AA.
XX
AC   Y70656;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature human transforming growth factor beta-9, Ztgfbeta-9 protein-3.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52195.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 85-86; 97pp; English.
XX
CC   The present sequence is the mature human transforming growth factor
CC   beta-9, designated Ztgf beta-9. This is a mature sequence excluding the
CC   signal sequence extending from amino acid 18 to and including amino acid
CC   202 of Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   185 AA;


  Query Match             96.6%;  Score 835;  DB 21;  Length 185;
  Best Local Similarity   98.7%;  Pred. No. 8.9e-86;
  Matches  154;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30 aagvlsafhhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyp 89

Qy      65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90 rylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsvyteayvtipvg 149

Qy     125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAGP 160
           ||||||||||||||||||||||||||||||||||||
Db     150 ctcvpepekdadsinssidkqgakllgpndapagp 185


RESULT   4
Y70655
ID   Y70655 standard; Protein; 186 AA.
XX
AC   Y70655;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature human transforming growth factor beta-9, Ztgfbeta-9 protein-2.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52195.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 84; 97pp; English.
XX
CC   The present sequence is the mature human transforming growth factor
CC   beta-9, designated Ztgf beta-9. This is a mature sequence excluding the
CC   signal sequence extending from amino acid 17 to and including amino acid
CC   202 of Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   186 AA;


  Query Match             96.6%;  Score 835;  DB 21;  Length 186;
  Best Local Similarity   98.7%;  Pred. No. 9e-86;
  Matches  154;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31 aagvlsafhhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyp 90

Qy      65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91 rylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsvyteayvtipvg 150

Qy     125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAGP 160
           ||||||||||||||||||||||||||||||||||||
Db     151 ctcvpepekdadsinssidkqgakllgpndapagp 186


RESULT   5
Y70654
ID   Y70654 standard; Protein; 187 AA.
XX
AC   Y70654;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature human transforming growth factor beta-9, Ztgfbeta-9 protein-1.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52195.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 84; 97pp; English.
XX
CC   The present sequence is the mature human transforming growth factor
CC   beta-9, designated Ztgf beta-9. This is a mature sequence excluding the
CC   signal sequence extending from amino acid 16 to and including amino acid
CC   202 of Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   187 AA;


  Query Match             96.6%;  Score 835;  DB 21;  Length 187;
  Best Local Similarity   98.7%;  Pred. No. 9.1e-86;
  Matches  154;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32 aagvlsafhhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyp 91

Qy      65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92 rylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsvyteayvtipvg 151

Qy     125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAGP 160
           ||||||||||||||||||||||||||||||||||||
Db     152 ctcvpepekdadsinssidkqgakllgpndapagp 187


RESULT   6
Y70663
ID   Y70663 standard; Protein; 187 AA.
XX
AC   Y70663;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature human Ztgfbeta-9 variant protein.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52201.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 93-94; 97pp; English.
XX
CC   The present sequence is a mature variant human transforming growth
CC   factor beta-9 protein, designated Ztgf beta-9. Human Ztgf beta-9 was
CC   isolated from an arrayed pituitary gland cDNA plasmid library by PCR
CC   screening. This can be used to treat a variety of neurodegenerative
CC   diseases such as amyotrophic lateral sclerosis (ALS), Alzheimer's
CC   disease, Huntington's disease, Parkinson's disease and peripheral
CC   neuropathies, or demyelinating diseases including multiple sclerosis.
CC   Ztgf beta-9 peptides have antiviral activity and may also be used to
CC   regulate the proliferation, differentiation and apoptosis of neurons,
CC   glial cells, lymphocytes, hematopoietic cells and stromal cells.
XX
SQ   Sequence   187 AA;


  Query Match             96.6%;  Score 835;  DB 21;  Length 187;
  Best Local Similarity   98.7%;  Pred. No. 9.1e-86;
  Matches  154;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32 aagvlsafhhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyp 91
```

```
Qy      65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92 rylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsvyteayvtipvg 151

Qy     125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAGP 160
           ||||||||||||||||||||||||||||||||||||
Db     152 ctcvpepekdadsinssidkqgakllgpndapagp 187


RESULT   7
Y70653
ID   Y70653 standard; Protein; 202 AA.
XX
AC   Y70653;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human transforming growth factor beta-9, Ztgfbeta-9 protein.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..15
FT                   /label= Signal_peptide
FT   Protein         16..202
FT                   /note= "Mature transforming growth factor beta-9"
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52195.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 83; 97pp; English.
XX
CC   The present sequence is the human transforming growth factor beta-9,
CC   designated Ztgf beta-9. Human Ztgf beta-9 was isolated from an arrayed
CC   pituitary gland cDNA plasmid library by PCR screening. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   202 AA;


  Query Match             96.6%;  Score 835;  DB 21;  Length 202;
  Best Local Similarity   98.7%;  Pred. No. 1e-85;
  Matches  154;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47 aagvlsafhhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyp 106

Qy      65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     107 rylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsvyteayvtipvg 166

Qy     125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAGP 160
           ||||||||||||||||||||||||||||||||||||
Db     167 ctcvpepekdadsinssidkqgakllgpndapagp 202


RESULT   8
Y70662
ID   Y70662 standard; Protein; 209 AA.
XX
AC   Y70662;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human transforming growth factor beta-9, Ztgfbeta-9 variant protein.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /label= Signal_peptide
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52201.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 92-93; 97pp; English.
XX
CC   The present sequence is a variant of human transforming growth
CC   factor beta-9, designated Ztgf beta-9. Human Ztgf beta-9 was isolated
CC   from an arrayed pituitary gland cDNA plasmid library by PCR screening.
CC   This can be used to treat a variety of neurodegenerative diseases such
CC   as amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's
CC   disease, Parkinson's disease and peripheral neuropathies, or
CC   demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides
CC   have antiviral activity and may also be used to regulate the
CC   proliferation, differentiation and apoptosis of neurons, glial cells,
CC   lymphocytes, hematopoietic cells and stromal cells.
XX
SQ   Sequence   209 AA;


  Query Match             96.6%;  Score 835;  DB 21;  Length 209;
  Best Local Similarity   98.7%;  Pred. No. 1e-85;
  Matches  154;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54 aagvlsafhhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyp 113

Qy      65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114 rylpeayclcrgcltglfgeedvrfrsapvymptvvlrrtpacaggrsvyteayvtipvg 173

Qy     125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAGP 160
           ||||||||||||||||||||||||||||||||||||
Db     174 ctcvpepekdadsinssidkqgakllLgpndapagp 209


RESULT   9
Y70658
ID   Y70658 standard; Protein; 183 AA.
XX
AC   Y70658;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Mature murine transforming growth factor beta-9, Ztgfbeta-9 protein.
XX
KW   Murine transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Mus sp.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52198.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 89; 97pp; English.
XX
CC   The present sequence is the mature murine transforming growth factor
CC   beta-9, designated Ztgf beta-9. The signal sequence extends from
CC   amino acid residues 1 to 22. Murine Ztgf beta-9 was found to be highly
CC   expressed in the HCL hypothalamic cell line. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   183 AA;


  Query Match             81.8%;  Score 707;  DB 21;  Length 183;
  Best Local Similarity   84.5%;  Pred. No. 1.8e-71;
  Matches  131;  Conservative    6;  Mismatches   14;  Indels    4;  Gaps    1;

Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
           |  |||||||||||||||||||||||| |||||||||||||||||||||||||||||:|

Db      32 aagvlsafhhtlqlgpreqarnascpaggraadrrfrpptnlrsvspwayrisydparfp 91

Qy      65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
           ||||||||||||||||:|||| |||| ||: | |||||| ||||||||| | |:||||
Db      92 rylpeayclcrgcltglygeedfrfrstpvfspavvlrrtaacaggrsvyaehyitipvg 151

Qy     125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAG 159
           |||||||:| ||| |||:|    |||||| | |||
Db     152 ctcvpepdksadsanssmd----kLllgpadrpag 182


RESULT  10
Y70657
ID   Y70657 standard; Protein; 205 AA.
XX
AC   Y70657;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Murine transforming growth factor beta-9, Ztgf beta-9 protein.
XX
KW   Murine transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /label= Signal_peptide
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
DR   N-PSDB; Z52198.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 88; 97pp; English.
XX
CC   The present sequence encodes murine transforming growth factor beta-9,
CC   designated Ztgf beta-9. Murine Ztgf beta-9 was found to be highly
CC   expressed in the HCL hypothalamic cell line. This can be used
CC   to treat a variety of neurodegenerative diseases such as amyotrophic
CC   lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC   Parkinson's disease and peripheral neuropathies, or demyelinating
CC   diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC   antiviral activity and may also be used to regulate the proliferation,
CC   differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC   hematopoietic cells and stromal cells.
XX
SQ   Sequence   205 AA;


  Query Match             81.8%;  Score 707;  DB 21;  Length 205;
  Best Local Similarity   84.5%;  Pred. No. 2.1e-71;
  Matches  131;  Conservative    6;  Mismatches   14;  Indels    4;  Gaps    1;

Qy       5 ARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYP 64
```

```
          |  |||||||||||||||||||||||||| |||||||||||||||||||||||||||:|
Db     54 aagvlsafhhtlqlgpreqarnascpaggraadrrfrpptnlrsvspwayrisydparfp 113

Qy     65 RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVG 124
          ||||||||||||||||:|||| |||| ||: | |||||| |||||||||| | |:||||||
Db    114 rylpeayclcrgcltglygeedfrfrstpvfspavvlrrtaacaggrsvyaehyitipvg 173

Qy    125 CTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPAG 159
          |||||||:| ||| |||:|    |||||| | |||
Db    174 ctcvpepdksadsanssmd----klllgpadrpag 204


RESULT 11
Y70666
ID   Y70666 standard; Protein; 57 AA.
XX
AC   Y70666;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human transforming growth factor beta-9, Ztgfbeta-9 peptide-6.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 94-95; 97pp; English.
XX
CC   The present sequence is an epitope-bearing peptide derived from the human
CC   transforming growth factor beta-9, designated Ztgf beta-9. These peptides
CC   are used to raise antibodies, including anti-idiotypic antibodies. This
CC   can be used to treat a variety of neurodegenerative diseases such as
CC   amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's
CC   disease, Parkinson's disease and peripheral neuropathies, or
CC   demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides
CC   have antiviral activity and may also be used to regulate the
CC   proliferation, differentiation and apoptosis of neurons, glial cells,
CC   lymphocytes, hematopoietic cells and stromal cells.
XX
SQ   Sequence   57 AA;


  Query Match             35.1%;  Score 303;  DB 21;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-27;
  Matches   57;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    102 RRTPACAGGRSVYTEAYVTIPVGCTCVPEPEKDADSINSSIDKQGAKLLLGPNDAPA 158
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 rrtpacaggrsvyteayvtipvgctcvpepekdadsinssidkqgakllgpndapa 57


RESULT 12
Y70664
ID   Y70664 standard; Protein; 54 AA.
XX
AC   Y70664;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human transforming growth factor beta-9, Ztgfbeta-9 peptide-4.
XX
KW   Human transforming growth factor beta-9; Ztgf beta-9;
KW   Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW   peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW   antiviral; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200015798-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US21677.
XX
PR   17-SEP-1998;   98US-0154817.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Presnell SR,  Taft DW,  Foley KP;
XX
DR   WPI; 2000-271436/23.
XX
PT   Polynucleotides encoding a novel transforming growth factor beta-9
PT   polypeptide, designated Ztgf beta-9, useful as an antiviral and
PT   antiproliferative agent  -
XX
PS   Claim 6; Page 94; 97pp; English.
XX
CC   The present sequence is an epitope-bearing peptide derived from the human
CC   transforming growth factor beta-9, designated Ztgf beta-9. These peptides
CC   are used to raise antibodies, including anti-idiotypic antibodies. This
CC   can be used to treat a variety of neurodegenerative diseases such as
CC   amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's
CC   disease, Parkinson's disease and peripheral neuropathies, or
CC   demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides
CC   have antiviral activity and may also be used to regulate the
CC   proliferation, differentiation and apoptosis of neurons, glial cells,
CC   lymphocytes, hematopoietic cells and stromal cells.
XX
SQ   Sequence   54 AA;


  Query Match             34.5%;  Score 298;  DB 21;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-26;
  Matches   54;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     17 QLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPRYLPEA 70
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 qlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyprylpea 54


RESULT 13
Y44462
ID   Y44462 standard; Protein; 425 AA.
XX
AC   Y44462;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Human Interleukin 17C-IgG1 Fc fusion protein, hIL-17C.fc.
XX
KW   Interleukin; IL-17C.fc; fusion protein; PRO1122 polypeptide; cytokine;
KW   human IgG1; fluoresence-activated cell sorter analysis; FACS;
```

```
KW   Tumour Necrosis Factor-alpha; TNF-alpha; leukemic monocyte; THP-1 cell.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /label= Signal_peptide
FT   Protein         19..197
FT                   /label= Mature_IL-17C_polypeptide
FT                   /note= "Used to treat degenerative cartilaginous
FT                   disorder"
FT   Misc-difference 109
FT                   /note= "Conserved Trp residue"
FT   Misc-difference 129
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 134
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 163
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 189
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 191
FT                   /note= "Conserved Cys residue"
FT   Region          197..425
FT                   /note= "Sequence derived from Fc region of human IgG1"
XX
PN   WO9960127-A2.
XX
PD   25-NOV-1999.
XX
PF   14-MAY-1999;   99WO-US10733.
XX
PR   15-MAY-1998;   98US-0085579.
PR   23-DEC-1998;   98US-0113621.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Filvaroff E,  Goddard A,  Gurney AL,  Li H,  Wood WI;
XX
DR   WPI; 2000-116314/10.
XX
PT   New polypeptides designated PRO1031 and PRO1122 used to treat a
PT   degenerative cartilaginous disorder -
XX
PS   Example 12; Page 129-130; 141pp; English.
XX
CC   The present sequence is the human IL-17C.fc fusion protein, derived from
CC   PRO1122 polypeptide and the Fc region of human IgG1. The cytokine IL-17C
CC   can be used to induce the release of TNF-alpha from human leukemic
CC   monocytic, THP-1 cells. The fusion protein, IL-17C.fc is used to identify
CC   the binding of IL-17C to THP-1 cells, using fluoresence-activated cell
CC   sorter analysis (FACS).
XX
SQ   Sequence   425 AA;

  Query Match             22.9%;  Score 197.5;  DB 21;  Length 425;
  Best Local Similarity   32.7%;  Pred. No. 5.3e-14;
  Matches   51;  Conservative   21;  Mismatches   73;  Indels   11;  Gaps    5;

Qy       7 AVLSAFHHTLQLGPREQ-ARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPR 65
           |::|:       |  |: :    ||   || :      |: ||:||| ||:  |  |||:
Db      65 alvssleaashrgrherpsattqcpv-lrp-eevleadthqrsispwryrvdtdedrypq 122

Qy      66 YLPEAYCLCRGCLTGLFGEEDVKFRSAPVYMPTVVLRRTPACAGGRSVYTEA-------Y 118
             |  | ||||||:    | |     |  :    :|||| |    |  : |        :
Db     123 klafaeclcrgcidartgretaalnsvrllqsllvlrrrpcsrdgsglptpgafafhtef 182

Qy     119 VTIPVGCTCVPEPEKDADSINSSIDKQGAKLLLGPN 154
           : :|||||||  |    |  ::       :|| ||:
Db     183 ihvpvgctcv-lprsvpdkthtcppcpapellggps 217
```

```
RESULT  14
Y92238
ID   Y92238 standard; Protein; 197 AA.
XX
AC   Y92238;
XX
DT   10-AUG-2000  (first entry)
XX
DE   Human interleukin-17 (IL-17) homologue.
XX
KW   Interleuk 17; IL-17; haematopoiesis; chemotherapy; cytostatic;
KW   antianemic; cardiant; hemostatic; anti-inflammatory; anti-HIV.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /label= signal_peptide
FT   Protein         19..197
FT                   /label= mature_protein
XX
PN   WO200020593-A1.
XX
PD   13-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22678.
XX
PR   02-OCT-1998;   98US-0102883.
PR   01-DEC-1998;   98US-0110405.
PR   11-JUN-1999;   99US-0138910.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Glasebrook AL,  Su EW,  Wei J,  Liu L;
XX
DR   WPI; 2000-303778/26.
DR   N-PSDB; A09153.
XX
PT   Nucleic acid encoding an interleukin-17 (IL-17) homolog polypeptide
PT   which enhances hematopoiesis, useful for treating e.g. anemia,
PT   thrombocytopenia, viral and bacterial infections
XX
PS   Claim 16; Page 92-93; 111pp; English.
XX
CC   Interleukin 17 (IL-17) stimulates hematopoiesis and production of
CC   neutrophils, granulocytes, or platelets, this may be useful during
CC   chemotherapy.  IL-17 homologues have at least one activity selected
CC   from induction of cytotoxic T cells, induction of lymphokine-activated
CC   killer cell proliferation or a B or T cell stimulation.  The IL-17
CC   homologue may also be used to treat viral or bacterial infections,
CC   immune related diseases, anemia, leukemia, thrombocytopenia, uremia,
CC   Von Willebrand disease,  postoperative cardiovascular dysfunction,
CC   treatment of AIDS (acquired immune deficiency syndrome)-related bone
CC   marrow failure, and inflammatory diseases of the gastrointestinal
CC   system, joints, and lungs.
XX
SQ   Sequence   197 AA;

  Query Match             22.2%;  Score 192;  DB 21;  Length 197;
  Best Local Similarity   34.6%;  Pred. No. 8.4e-14;
  Matches   45;  Conservative   17;  Mismatches   58;  Indels   10;  Gaps    4;

Qy       7 AVLSAFHHTLQLGPREQ-ARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPR 65
           |::|:       |  |: :    ||   || :      |: ||:||| ||:  |  |||:
Db      65 alvssleaashrgrherpsattqcpv-lrp-eevleadthqrsispwryrvdtdedrypq 122

Qy      66 YLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEA-------Y 118
             |  | ||||||:    | |     |  :    :|||| |    |  : |        :
Db     123 klafaeclcrgcidartgretaalnsvrllqsllvlrrrpcsrdgsglptpgafafhtef 182

Qy     119 VTIPVGCTCV 128
           : :|||||||
```

```
Db    183 ihvpvgctcv 192


RESULT  15
Y44460
ID   Y44460 standard; Protein; 197 AA.
XX
AC   Y44460;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Human Interleukin 17C, PRO1122 polypeptide.
XX
KW   Interleukin; IL-17C; PRO1122 polypeptide; clone DNA62377-1381-1; UNQ561;
KW   cytokine IL-17; cytotoxic T-lymphocyte-associated antigen 8; CTLA-8;
KW   hybridisation probe; antagonist; degenerative cartilaginous disorder;
KW   agonist; diagnose; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /label= Signal_peptide
FT   Protein         19..197
FT                   /label= Mature_IL-17C_polypeptide
FT                   /note= "Used to treat degenerative cartilaginous
FT                   disorder"
FT   Misc-difference 109
FT                   /note= "Conserved Trp residue"
FT   Misc-difference 129
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 134
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 163
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 189
FT                   /note= "Conserved Cys residue"
FT   Misc-difference 191
FT                   /note= "Conserved Cys residue"
XX
PN   WO9960127-A2.
XX
PD   25-NOV-1999.
XX
PF   14-MAY-1999;   99WO-US10733.
XX
PR   15-MAY-1998;   98US-0085579.
PR   23-DEC-1998;   98US-0113621.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Filvaroff E,  Goddard A,  Gurney AL,  Li H,  Wood WI;
XX
DR   WPI; 2000-116314/10.
DR   N-PSDB; Z29728.
XX
PT   New polypeptides designated PRO1031 and PRO1122 used to treat a
PT   degenerative cartilaginous disorder -
XX
PS   Claim 23; Fig 3; 141pp; English.
XX
CC   The present sequence is the human PRO1122 polypeptide, also referred to
CC   as UNQ561, and as interleukin-17C (IL-17C), encoded by
CC   clone DNA62377-1381-1. This sequence has identity with the
CC   cytokine IL-17 and cytotoxic T-lymphocyte-associated antigen 8 (CTLA-8)
CC   and has leucine zipper pattern. PRO1122 is expressed in pancreas, small
CC   intestine, stomach and testis also. It shares about 26-28% amino acid
CC   identity with IL-17 and IL-17B. The entire coding region of IL-17C can
CC   be used as hybridisation probe. The PRO1122 polypeptide, agonist or
CC   antagonist, is used to diagnose and treat a degenerative cartilaginous
CC   disorder.
XX
SQ   Sequence   197 AA;


  Query Match             22.2%;  Score 192;  DB 21;  Length 197;
  Best Local Similarity   34.6%;  Pred. No. 8.4e-14;
  Matches   45;  Conservative   17;  Mismatches   58;  Indels   10;  Gaps    4;

Qy       7 AVLSAFHHTLQLGPREQ-ARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPR 65
           |::|:       |  |: :    ||   || :      |: ||:||| ||:  |  |||:
Db      65 alvssleaashrgrherpsattqcpv-lrp-eevleadthqrsispwryrvdtdedrypq 122

Qy      66 YLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEA-------Y 118
            |  | ||||||:    | |     |  :    :|||| |    |  : |        :
Db     123 klafaeclcrgcidartgretaalnsvrllqsllvlrrrpcsrdgsglptpgafafhtef 182

Qy     119 VTIPVGCTCV 128
           : :||||||| 
Db     183 ihvpvgctcv 192


Search completed: February 16, 2001, 12:49:48
Job time: 113 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:50:10 ; Search time 18.64 Seconds
(without alignments)
154.138 Million cell updates/sec

Title: US-09-320-713-4
Perfect score: 864
Sequence: 1 NSARARAVLSAFHHTLQLGP.....SIDKQGAKLLLGPNDAPAGP 160

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	159	18.4	153	1	US-08-514-014-12	Sequence 12, Appl
2	159	18.4	153	2	US-08-833-823-12	Sequence 12, Appl
3	159	18.4	163	3	US-09-034-810-2	Sequence 2, Appli
4	159	18.4	163	3	US-08-685-239-2	Sequence 2, Appli
5	112	13.0	158	2	US-08-620-694A-7	Sequence 7, Appli
6	112	13.0	158	3	US-09-022-255-7	Sequence 7, Appli
7	112	13.0	158	3	US-09-022-696-7	Sequence 7, Appli
8	112	13.0	158	3	US-09-022-253-7	Sequence 7, Appli
9	112	13.0	158	3	US-09-022-260-7	Sequence 7, Appli
10	108.5	12.6	151	2	US-08-620-694A-8	Sequence 8, Appli
11	108.5	12.6	151	3	US-09-034-810-6	Sequence 6, Appli
12	108.5	12.6	151	3	US-09-022-255-8	Sequence 8, Appli
13	108.5	12.6	151	3	US-09-022-696-8	Sequence 8, Appli
14	108.5	12.6	151	3	US-08-685-239-6	Sequence 6, Appli
15	108.5	12.6	151	3	US-09-022-253-8	Sequence 8, Appli
16	108.5	12.6	151	3	US-09-022-260-8	Sequence 8, Appli
17	108	12.5	150	3	US-09-034-810-4	Sequence 4, Appli
18	108	12.5	150	3	US-08-685-239-4	Sequence 4, Appli
19	77.5	9.0	164	1	US-08-154-915-6	Sequence 6, Appli
20	75.5	8.7	397	4	PCT-US94-09700-11	Sequence 11, Appl
21	75	8.7	361	2	US-08-483-926A-9	Sequence 9, Appli
22	71.5	8.3	164	1	US-07-970-462A-2	Sequence 2, Appli
23	71.5	8.3	164	1	US-08-406-248-2	Sequence 2, Appli
24	71.5	8.3	164	1	US-08-574-043A-2	Sequence 2, Appli
25	71.5	8.3	164	2	US-08-524-218A-2	Sequence 2, Appli
26	71.5	8.3	164	2	US-08-795-015-2	Sequence 2, Appli
27	71.5	8.3	164	4	PCT-US92-10904-2	Sequence 2, Appli
28	71.5	8.3	164	4	PCT-US94-09700-2	Sequence 2, Appli
29	71.5	8.3	164	4	PCT-US94-12936-2	Sequence 2, Appli
30	71.5	8.3	164	4	PCT-US96-11886A-2	Sequence 2, Appli
31	70.5	8.2	120	5	5514582-36	Patent No. 5514582
32	70.5	8.2	164	1	US-08-275-983B-4	Sequence 4, Appli
33	70	8.1	299	3	US-09-188-930-192	Sequence 192, App
34	70	8.1	299	3	US-09-188-930-332	Sequence 332, App
35	69	8.0	362	2	US-08-737-045-13	Sequence 13, Appl
36	68	7.9	2205	1	US-08-093-453B-2	Sequence 2, Appli
37	67	7.8	199	3	US-09-188-930-290	Sequence 290, App
38	66	7.6	676	1	US-08-343-785-2	Sequence 2, Appli
39	66	7.6	676	1	US-08-343-785-8	Sequence 8, Appli
40	66	7.6	676	2	US-08-462-221-2	Sequence 2, Appli
41	66	7.6	676	2	US-08-462-221-8	Sequence 8, Appli
42	66	7.6	676	3	US-08-946-458-2	Sequence 2, Appli
43	66	7.6	676	3	US-08-946-458-8	Sequence 8, Appli
44	65	7.5	149	1	US-08-469-427A-14	Sequence 14, Appl
45	65	7.5	149	2	US-08-039-297B-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-514-014-12
; Sequence 12, Application US/08514014
; Patent No. 5707829
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    TITLE OF INVENTION:  DNA SEQUENCES AND SECRETED PROTEINS
;    TITLE OF INVENTION:  ENCODED THEREBY
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc. -- Legal Affairs
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/514,014
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI6000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  153 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-514-014-12


  Query Match           18.4%;  Score 159;  DB 1;  Length 153;
  Best Local Similarity  32.6%;  Pred. No. 6.4e-11;
  Matches   46;  Conservative   18;  Mismatches   53;  Indels   24;  Gaps    7;

Qy       1 NSARARAVLSAFHHTLQLGPREQARNASC---PAGGRPAD-------RRFRPPTNL--RS 48
           : | || :    |   |       :  ||   | |    |       :|     |:  ||
Db      16 SEAAARKIPKVGHTFFQ-------KPESCPPVPGGSMKLDIGIINENQRVSMSRNIESRS 68
```

```
Qy      49 VSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACA 108
           ||| | :::|| ||| : :| |   ||:    |:||:   | |:   |:|:||
Db      69 TSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQ-- 125

Qy     109 GGRSV-YTEAYVTIPVGCTCV 128
           | || :    | : ||||||
Db     126 -GCSVSFQLEKVLVTVGCTCV 145


RESULT   2
US-08-833-823-12
; Sequence 12, Application US/08833823
; Patent No. 5969093
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    TITLE OF INVENTION:  DNA SEQUENCES AND SECRETED PROTEINS
;    TITLE OF INVENTION:  ENCODED THEREBY
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc. -- Legal Affairs
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/833,823
;      FILING DATE:  10-APR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/514,014
;      FILING DATE:  11-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI6000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  153 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-833-823-12


  Query Match             18.4%;  Score 159;  DB 2;  Length 153;
  Best Local Similarity   32.6%;  Pred. No. 6.4e-11;
  Matches   46;  Conservative   18;  Mismatches   53;  Indels   24;  Gaps    7;

Qy       1 NSARARAVLSAFHHTLQLGPREQARNASC---PAGGRPAD-------RRFRPPTNL--RS 48
           : | || :    |   |       :  ||   | |    |       :|     |:  ||
Db      16 SEAAARKIPKVGHTFFQ-------KPESCPPVPGGSMKLDIGIINENQRVSMSRNIESRS 68

Qy      49 VSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACA 108
           ||| | :::|| ||| : :| |   ||:    |:||:   | |:   |:|:||
Db      69 TSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQ-- 125

Qy     109 GGRSV-YTEAYVTIPVGCTCV 128
           | || :    | : ||||||
Db     126 -GCSVSFQLEKVLVTVGCTCV 145


RESULT   3
US-09-034-810-2
; Sequence 2, Application US/09034810
; Patent No. 6043344
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/034,810
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/685,239
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-034-810-2


  Query Match             18.4%;  Score 159;  DB 3;  Length 163;
  Best Local Similarity   32.6%;  Pred. No. 7e-11;
  Matches   46;  Conservative   18;  Mismatches   53;  Indels   24;  Gaps    7;

Qy       1 NSARARAVLSAFHHTLQLGPREQARNASC---PAGGRPAD-------RRFRPPTNL--RS 48
           : | || :    |   |       :  ||   | |    |       :|     |:  ||
Db      26 SEAAARKIPKVGHTFFQ-------KPESCPPVPGGSMKLDIGIINENQRVSMSRNIESRS 78

Qy      49 VSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACA 108
           ||| | :::|| ||| : :| |   ||:    |:||:   | |:   |:|:||
Db      79 TSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQ-- 135

Qy     109 GGRSV-YTEAYVTIPVGCTCV 128
           | || :    | : ||||||
Db     136 -GCSVSFQLEKVLVTVGCTCV 155


RESULT   4
US-08-685-239-2
; Sequence 2, Application US/08685239
; Patent No. 6074849
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,239
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  163 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-685-239-2


  Query Match             18.4%;  Score 159;  DB 3;  Length 163;
  Best Local Similarity   32.6%;  Pred. No. 7e-11;
  Matches   46;  Conservative   18;  Mismatches   53;  Indels   24;  Gaps    7;

Qy       1 NSARARAVLSAFHHTLQLGPREQARNASC---PAGGRPAD-------RRFRPPTNL--RS 48
           : | || :    |   |       :  ||   | |    |       :|     |:  ||
Db      26 SEAAARKIPKVGHTFFQ-------KPESCPPVPGGSMKLDIGIINENQRVSMSRNIESRS 78

Qy      49 VSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACA 108
           ||| | :::|| |||  : :| |   ||:    |:||:   | |:   |:|:||
Db      79 TSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQ-- 135

Qy     109 GGRSV-YTEAYVTIPVGCTCV 128
           | || :    | : ||||||
Db     136 -GCSVSFQLEKVLVTVGCTCV 155


RESULT   5
US-08-620-694A-7
; Sequence 7, Application US/08620694A
; Patent No. 5869286
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 5869286el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/620,694A
;      FILING DATE:  21 MARCH 1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/538,765
;      FILING DATE:  7 AUGUST 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH:  158 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-620-694A-7


  Query Match             13.0%;  Score 112;  DB 2;  Length 158;
  Best Local Similarity   33.3%;  Pred. No. 1.7e-05;
  Matches   33;  Conservative   14;  Mismatches   44;  Indels    8;  Gaps    5;

Qy      32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFR 90
           | : : |  ||    | || |||    : || |||  : || |  : |:    |: |
Db      58 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMN 114

Qy      91 SAPVYMPTVVLRRTP-ACAGGRSVYTEAYVTIPVGCTCV 128
           |  :    :||:| | :|      :    : : ||||||
Db     115 SVLIQQEILVLKREPESCP---FTFRVEKMLVGVGCTCV 150


RESULT   6
US-09-022-255-7
; Sequence 7, Application US/09022255
; Patent No. 6072033
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6072033el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,255
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/620,694
;      FILING DATE:  21 MARCH 1996
;      APPLICATION NUMBER:  USSN 08/538,765
;      FILING DATE:  7 AUGUST 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 158 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-022-255-7


  Query Match             13.0%;  Score 112;  DB 3;  Length 158;
  Best Local Similarity   33.3%;  Pred. No. 1.7e-05;
  Matches   33;  Conservative   14;  Mismatches   44;  Indels    8;  Gaps    5;

Qy      32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFR 90
           | : : |  ||   | || |||    : || |||  : || |  : |:    |: |
Db      58 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMN 114

Qy      91 SAPVYMPTVVLRRTP-ACAGGRSVYTEAYVTIPVGCTCV 128
           |  :    :||:| | :|      :    : : ||||||
Db     115 SVLIQQEILVLKREPESCP---FTFRVEKMLVGVGCTCV 150


RESULT   7
US-09-022-696-7
; Sequence 7, Application US/09022696
; Patent No. 6072037
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6072037el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,696
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/620,694
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 158 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-022-696-7


  Query Match             13.0%;  Score 112;  DB 3;  Length 158;
  Best Local Similarity   33.3%;  Pred. No. 1.7e-05;
  Matches   33;  Conservative   14;  Mismatches   44;  Indels    8;  Gaps    5;

Qy      32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFR 90
           | : : |  ||   | || |||    : || |||  : || |  : |:    |: |
Db      58 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMN 114

Qy      91 SAPVYMPTVVLRRTP-ACAGGRSVYTEAYVTIPVGCTCV 128
           |  :    :||:| | :|      :    : : ||||||
Db     115 SVLIQQEILVLKREPESCP---FTFRVEKMLVGVGCTCV 150


RESULT   8
US-09-022-253-7
; Sequence 7, Application US/09022253
; Patent No. 6096305
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6096305el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,253
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/620,694
;      FILING DATE:  21-MARCH-1996
;      APPLICATION NUMBER:  USSN 08/538,765
;      FILING DATE:  7 AUGUST 1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
```

```
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  158 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-022-253-7


  Query Match             13.0%;  Score 112;  DB 3;  Length 158;
  Best Local Similarity   33.3%;  Pred. No. 1.7e-05;
  Matches   33;  Conservative   14;  Mismatches   44;  Indels    8;  Gaps    5;

Qy      32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFR 90
           | : : |  ||   | || |||    : || |||  : || |  : |:    |: |
Db      58 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMN 114

Qy      91 SAPVYMPTVVLRRTP-ACAGGRSVYTEAYVTIPVGCTCV 128
           |  :    :||:| | :|      :    : : ||||||
Db     115 SVLIQQEILVLKREPESCP---FTFRVEKMLVGVGCTCV 150


RESULT  9
US-09-022-260-7
; Sequence 7, Application US/09022260
; Patent No. 6100235
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6100235el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,260
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/620,694
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 158 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-022-260-7


  Query Match             13.0%;  Score 112;  DB 3;  Length 158;
  Best Local Similarity   33.3%;  Pred. No. 1.7e-05;
  Matches   33;  Conservative   14;  Mismatches   44;  Indels    8;  Gaps    5;

Qy      32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFR 90
           | : : |  ||   | || |||    : || |||  : || |  : |:    |: |
Db      58 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMN 114

Qy      91 SAPVYMPTVVLRRTP-ACAGGRSVYTEAYVTIPVGCTCV 128
           |  :    :||:| | :|      :    : : ||||||
Db     115 SVLIQQEILVLKREPESCP---FTFRVEKMLVGVGCTCV 150


RESULT  10
US-08-620-694A-8
; Sequence 8, Application US/08620694A
; Patent No. 5869286
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 5869286el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/620,694A
;      FILING DATE:  21 MARCH 1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/538,765
;      FILING DATE:  7 AUGUST 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpesvirus Saimiri
;      STRAIN:  ORF13
US-08-620-694A-8


  Query Match             12.6%;  Score 108.5;  DB 2;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 4.1e-05;
  Matches   27;  Conservative   11;  Mismatches   37;  Indels   11;  Gaps    3;
```

```
Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLHRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
             |      ::    : : ||||||
Db     121 ---GHQPCPNSFRLEKMLVTVGCTCV 143


RESULT 11
US-09-034-810-6
; Sequence 6, Application US/09034810
; Patent No. 6043344
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/034,810
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/685,239
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-034-810-6


  Query Match             12.6%;  Score 108.5;  DB 3;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 4.1e-05;
  Matches   27;  Conservative   11;  Mismatches   37;  Indels   11;  Gaps    3;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLHRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
             |      ::    : : ||.|||
Db     121 ---GHQPCPNSFRLEKMLVTVGCTCV 143


RESULT 12
US-09-022-255-8
; Sequence 8, Application US/09022255
; Patent No. 6072033
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6072033el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,255
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/620,694
;      FILING DATE:  21 MARCH 1996
;      APPLICATION NUMBER:  USSN 08/538,765
;      FILING DATE:  7 AUGUST 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpesvirus Saimiri
;      STRAIN:  ORF13
US-09-022-255-8


  Query Match             12.6%;  Score 108.5;  DB 3;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 4.1e-05;
  Matches   27;  Conservative   11;  Mismatches   37;  Indels   11;  Gaps    3;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLHRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
             |      ::    : : ||||||
Db     121 ---GHQPCPNSFRLEKMLVTVGCTCV 143


RESULT 13
US-09-022-696-8
; Sequence 8, Application US/09022696
; Patent No. 6072037
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6072037el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
;      OPERATING SYSTEM:  Apple Operating System 7.5.5
;      SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/022,696
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/620,694
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/410,535
;      FILING DATE:  23 MARCH 1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perkins, Patricia Anne
;      REGISTRATION NUMBER:  34,695
;      REFERENCE/DOCKET NUMBER:  2617-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)587-0430
;      TELEFAX:  (206)
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpesvirus Saimiri
;      STRAIN:  ORF13
US-09-022-696-8


  Query Match             12.6%;  Score 108.5;  DB 3;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 4.1e-05;
  Matches   27;  Conservative   11;  Mismatches   37;  Indels   11;  Gaps    3;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLHRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
             |      ::    : : ||||||
Db     121 ---GHQPCPNSFRLEKMLVTVGCTCV 143


RESULT  14
US-08-685-239-6
; Sequence 6, Application US/08685239
; Patent No. 6074849
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  Kelleher, Kerry
;    APPLICANT:  Carlin, McKeough
;    APPLICANT:  Goldman, Samuel
;    APPLICANT:  Pittman, Debra
;    APPLICANT:  Mi, Sha
;    APPLICANT:  Neben, Steven
;    APPLICANT:  Giannotti, JoAnn
;    APPLICANT:  Golden'Fleet, Margaret
;    TITLE OF INVENTION:  Human CTLA-8 and Uses of CTLA-8-Related Proteins
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,239
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  151 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-685-239-6


  Query Match             12.6%;  Score 108.5;  DB 3;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 4.1e-05;
  Matches   27;  Conservative   11;  Mismatches   37;  Indels   11;  Gaps    3;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLHRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
             |      ::    : : ||||||
Db     121 ---GHQPCPNSFRLEKMLVTVGCTCV 143


RESULT  15
US-09-022-253-8
; Sequence 8, Application US/09022253
; Patent No. 6096305
;  GENERAL INFORMATION:
;    APPLICANT:  Yao, Zhengbin
;    APPLICANT:  Spriggs, Melanie
;    APPLICANT:  Fanslow, William
;    TITLE OF INVENTION:  No. 6096305el Receptor That Binds IL-17
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Power Macintosh
```

```
;       OPERATING SYSTEM:  Apple Operating System 7.5.5
;       SOFTWARE:  Microsoft Word for Apple, Version 6.0.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/022,253
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/620,694
;       FILING DATE:  21-MARCH-1996
;       APPLICATION NUMBER:  USSN 08/538,765
;       FILING DATE:  7 AUGUST 1995
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  USSN 08/410,535
;       FILING DATE:  23 MARCH 1995
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Perkins, Patricia Anne
;       REGISTRATION NUMBER:  34,695
;       REFERENCE/DOCKET NUMBER:  2617-B
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206)587-0430
;       TELEFAX:  (206)
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  151 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  No. 6096305 Relevant
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Herpesvirus Saimiri
;       STRAIN:  ORF13
US-09-022-253-8


  Query Match             12.6%;  Score 108.5;  DB 3;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 4.1e-05;
  Matches   27;  Conservative   11;  Mismatches   37;  Indels   11;  Gaps    3;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLHRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
               |      ::    : : ||||||
Db     121 ---GHQPCPNSFRLEKMLVTVGCTCV 143


Search completed: February 16, 2001, 12:50:11
Job time: 131 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:50:39 ; Search time 23.57 Seconds
(without alignments)
460.929 Million cell updates/sec

Title: US-09-320-713-4
Perfect score: 864
Sequence: 1 NSARARAVLSAFHHTLQLGP.....SIDKQGAKLLLGPNDAPAGP 160

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	113.5	13.1	147	2	JC4628	cytotoxic T-lympho
2	108.5	12.6	151	1	B45351	immediate-early pr
3	108	12.5	150	2	I49623	cytotoxic T-lympho
4	99	11.5	148	2	T21334	hypothetical prote
5	83.5	9.7	226	2	T27843	hypothetical prote
6	77	8.9	354	2	T27099	hypothetical prote
7	77	8.9	603	2	S28941	coagulation factor
8	76	8.8	242	2	B81719	conserved hypothet
9	75.5	8.7	610	2	T44254	thiamin biosynthes
10	75	8.7	178	2	D72281	hypothetical prote
11	75	8.7	242	2	A71568	hypothetical prote
12	75	8.7	362	2	S22395	fetuin precursor -
13	74.5	8.6	181	2	I68674	cyclin-dependent k
14	74.5	8.6	1273	2	T38292	hypothetical prote
15	74.5	8.6	1896	1	RNFF2L	DNA-directed RNA p
16	74	8.6	495	2	T09993	diaminopimelate de
17	73	8.4	2397	1	A55535	versican precursor
18	71.5	8.3	164	2	A49437	cyclin-dependent k
19	71.5	8.3	789	2	A39564	transcription repr
20	71	8.2	464	2	D72653	hypothetical prote
21	71	8.2	998	2	G83022	probable two-compo
22	70.5	8.2	181	2	I54380	cyclin-dependent k
23	70.5	8.2	256	2	T47860	transcription fact
24	70.5	8.2	265	2	D70778	hypothetical prote
25	70.5	8.2	479	1	A42241	glycine hydroxymet
26	70.5	8.2	872	2	H75564	probable ATP-depen
27	70.5	8.2	2115	2	S38480	nonstructural prot
28	70	8.1	206	2	T36643	probable integral
29	70	8.1	816	2	C69493	hypothetical prote
30	69	8.0	162	2	T32515	hypothetical prote
31	69	8.0	292	2	T03122	hypothetical prote
32	69	8.0	802	2	T24293	hypothetical prote
33	69	8.0	949	2	T24294	hypothetical prote
34	68.5	7.9	262	2	F75093	hypothetical prote
35	68.5	7.9	342	2	A56552	homeotic protein H
36	68	7.9	173	2	E71017	probable NADH-ubiq
37	68	7.9	298	2	A41230	DNA-3-methyladenin
38	68	7.9	528	2	T00951	probable 3-oxoacyl
39	68	7.9	585	2	A83020	probable carbamoyl
40	68	7.9	990	2	T14756	hypothetical prote
41	68	7.9	2205	1	MNWVRN	nonstructural poly
42	67.5	7.8	298	2	A32872	myogenic factor CM
43	67.5	7.8	298	2	A53783	transcription regu
44	67.5	7.8	426	2	A42360	cellulase (EC 3.2.
45	67.5	7.8	453	2	T01114	hypothetical prote

ALIGNMENTS

RESULT 1
JC4628
cytotoxic T-lymphocyte-associated antigen 8 precursor - mouse
N;Alternate names: CTLA8 protein
C;Species: Mus musculus (house mouse)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 26-Aug-1999
C;Accession: JC4628
R;Yao, Z.; Timour, M.; Painter, S.; Fanslow, W.; Spriggs, M.
Gene 168, 223-225, 1996
A;Title: Complete nucleotide sequence of the mouse CTLA8 gene.
A;Reference number: JC4628; MUID:96194901
A;Accession: JC4628
A;Molecule type: DNA
A;Residues: 1-147 <YAO>
A;Cross-references: GB:U35108; NID:g1244499; PIDN:AAA93253.1; PID:g1244500
C;Genetics:
A;Gene: ctla8
A;Introns: 69/2
C;Superfamily: saimiri herpesvirus immediate-early protein 2
C;Keywords: cytokine; glycoprotein; lymphocyte
F;1-14/Domain: signal sequence #status predicted <SIG>
F;15-147/Product: cytotoxic T-lymphocyte-associated antigen 8 #status predicted <MAT>
F;60/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.1%; Score 113.5; DB 2; Length 147;
Best Local Similarity 31.4%; Pred. No. 0.0003;
Matches 37; Conservative 16; Mismatches 46; Indels 19; Gaps 6;

```
Qy      32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFR 90
           | : : |  ||   | || |||    : || |||  : || |  : |:    |: |
Db      47 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMN 103

Qy      91 SAPVYMPTVVLRRTP-ACAGGRSVYTEAYVTIPVGCTCVPEPEKDADSINSSIDKQGA 147
           |  :    :||:| | :|      :    : : ||||||           :|| :| |
Db     104 SVLIQQEILVLKREPESCP---FTFRVEKMLVGVGCTCV----------ASIVRQAA 147
```

RESULT 2
B45351
immediate-early protein 2 - saimiriine herpesvirus 1 (strain 11)
N;Alternate names: hypothetical protein ORF13
C;Species: saimiriine herpesvirus 1
A;Note: host Saimiri sciureus (common squirrel monkey)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C;Accession: B45351; D36807
R;Nicholas, J.; Smith, E.P.; Coles, L.; Honess, R.
Virology 179, 189-200, 1990
A;Title: Gene expression in cells infected with gammaherpesvirus saimiri: properties
A;Reference number: A45351; MUID:91021021
A;Accession: B45351
A;Molecule type: mRNA

```
A;Residues: 1-151 <NIC>
A;Cross-references: GB:M60286; NID:g331040; PIDN:AAA46156.1; PID:g331042
R;Albrecht, J.
submitted to the EMBL Data Library, January 1992
A;Description: Primary structure of the herpesvirus saimiri genome.
A;Reference number: A36806
A;Accession: D36807
A;Molecule type: DNA
A;Residues: 1-151 <ALB>
A;Cross-references: GB:X64346; NID:g60320; PIDN:CAA45636.1; PID:g60334
R;Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; Wi
J. Virol. 66, 5047-5058, 1992
A;Title: Primary structure of the herpesvirus saimiri genome.
A;Reference number: A37309; MUID:92333688
A;Contents: annotation; protein-coding frames
A;Note: neither protein nor nucleotide sequence is given
C;Genetics:
A;Gene: 13
C;Superfamily: saimiri herpesvirus immediate-early protein 2
C;Keywords: early protein


  Query Match             12.6%;  Score 108.5;  DB 1;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 0.00096;
  Matches   27;  Conservative   11;  Mismatches   37;  Indels   11;  Gaps    3;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLHRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
           |      ::    : : || |||
Db     121 ---GHQPCPNSFRLEKMLVTVGCTCV 143


RESULT   3
I49623
cytotoxic T-lymphocyte-associated antigen 8 precursor - mouse
N;Alternate names: immediate-early protein 2 (ORF13) homolog
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 26-Aug-1999
C;Accession: I49623
R;Rouvier, E.; Luciani, M.
J. Immunol. 150, 5445-5456, 1993
A;Title: CTLA-8, cloned from an activated T cell, bearing AU-rich messenger RNA instabil
A;Reference number: I49623; MUID:93294300
A;Accession: I49623
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-150 <RES>
A;Cross-references: GB:L13839; NID:g293329; PIDN:AAA37490.1; PID:g293330
C;Genetics:
A;Gene: Ctla-8
C;Superfamily: saimiri herpesvirus immediate-early protein 2


  Query Match             12.5%;  Score 108;  DB 2;  Length 150;
  Best Local Similarity   34.4%;  Pred. No. 0.0011;
  Matches   31;  Conservative   11;  Mismatches   42;  Indels    6;  Gaps    4;

Qy      41 RPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTV 99
           ||    | || |||    : || |||  : || |  : |:    |: |    |  :    :
Db      57 RPSDYLNRSTSPWTLSRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMNSVLIQQEIL 115

Qy     100 VLRRTP-ACAGGRSVYTEAYVTIPVGCTCV 128
           ||:| |  |      :    : : ||||||
Db     116 VLKREPEKCP---FTFRVEKMLVGVGCTCV 142


RESULT   4
T21334
hypothetical protein F25D1.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21334
R;Kelly, P.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19408
A;Accession: T21334
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-148 <WIL>
A;Cross-references: EMBL:Z73973; PIDN:CAA98268.1; GSPDB:GN00023; CESP:F25D1.3
A;Experimental source: clone F25D1
C;Genetics:
A;Gene: CESP:F25D1.3
A;Map position: 5
A;Introns: 60/2


  Query Match             11.5%;  Score 99;  DB 2;  Length 148;
  Best Local Similarity   29.8%;  Pred. No. 0.0085;
  Matches   28;  Conservative   17;  Mismatches   35;  Indels   14;  Gaps    5;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGC--LTGLFGEEDVRFRSAPVYMPTVVLRRT 104
           |:: ||  |::|  :| |: : |: ||||     || |          |:     :|||
Db      52 RALCPWDSRVNYQESREPKLIAESVCLCRKSRGSTGAF--------CMPIVRKVPILRRV 103

Qy     105 PAC--AGGRSVYTEAYVTIPVGC-TCVPEPEKDA 135
           :|  : |   |  :   | ||| : :|  :: |
Db     104 -SCDRSTGLWNYVRSTELITVGCHSVLPRTQRAA 136


RESULT   5
T27843
hypothetical protein ZK39.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27843
R;Kershaw, J.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z20428
A;Accession: T27843
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-226 <WIL>
A;Cross-references: EMBL:Z82093; PIDN:CAB05021.1; GSPDB:GN00019; CESP:ZK39.
A;Experimental source: clone ZK39
C;Genetics:
A;Gene: CESP:ZK39.8
A;Map position: 1
A;Introns: 32/1; 52/3; 108/1


  Query Match             9.7%;  Score 83.5;  DB 2;  Length 226;
  Best Local Similarity   25.6%;  Pred. No. 0.48;
  Matches   32;  Conservative   18;  Mismatches   34;  Indels   41;  Gaps    8;

Qy      26 NASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPRYLPEAYCLCRGC-LTGLFGE 84
           |  | ||      ||  |:       |  |:   |  | : | |: |  :|  |||:  :
Db      52 NRGCDAGW----TRFNRPS-----GGWCVRVF--PGTYHQPLAESRCQSQGAVLTGVQNQ 100

Qy      85 EDVRFRSAPVYMPTVV---------LRRTPACAGGRSVYTEAYVTIPVGCTCVPEPEKDA 135
           |: : : | : :| :          | |||||:           |:  :|
Db     101 EEAK-KIASLLLPQISQQSGSIYIGLHRTPACSKS----------PISSSC-------- 140

Qy     136 DSINS 140
           :|:||
Db     141 NSMNS 145


RESULT   6
T27099
hypothetical protein Y51H1A.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
```

```
C;Accession: T27099
R;Smye, R.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20309
A;Accession: T27099
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-354 <WIL>
A;Cross-references: EMBL:AL032644; PIDN:CAA21670.1; GSPDB:GN00020; CESP:Y51H1A.7
A;Experimental source: clone Y51H1A
C;Genetics:
A;Gene: CESP:Y51H1A.7
A;Map position: 2
A;Introns: 310/3


  Query Match             8.9%;  Score 77;  DB 2;  Length 354;
  Best Local Similarity   25.0%;  Pred. No. 3.5;
  Matches   36;  Conservative   20;  Mismatches   56;  Indels   32;  Gaps    8;

Qy      20 PREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARY-PRYLPEAYCLCRGCL 78
           |  : | :: ||   ||  ::||    ||   :  ||      | ||  |:|| | :
Db     141 PERRPRRSAAPAS--PAKPQYRPQPPARSHEEYTQRIMNKVKEYMPRRRPDAYDLQQ--- 195

Qy      79 TGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVGCTCVPEPEKDADSI 138
               |::  | :     :    | | |       :  |     |   |  | | |::
Db     196 ----GQQQSRLQ----MLREQTLLRAP-------LEEEK----PPVATMAPVPDE----- 231

Qy     139 NSSI--DKQGAKLLLGPNDAPAGP 160
           : :|   : | |  :|| :|   |
Db     232 HKAIPWTRPGRKRAMGPGEAATAP 255


RESULT   7
S28941
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N;Alternate names: Hageman factor
C;Species: Cavia porcellus (guinea pig)
C;Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C;Accession: S28941
R;Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe, H.
Biochim. Biophys. Acta 1159, 113-121, 1992
A;Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage si
A;Reference number: S28941; MUID:93003367
A;Accession: S28941
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-603 <SEM>
A;Cross-references: EMBL:X68615; NID:g49578; PIDN:CAA48600.1; PID:g49579
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C;Keywords: hydrolase; serine proteinase
F;46-87/Domain: fibronectin type II repeat homology <1F2>
F;134-169/Domain: fibronectin type I repeat homology <FB1>
F;177-208/Domain: EGF homology <EGF>
F;216-294/Domain: kringle homology <KRG>
F;359-597/Domain: trypsin homology <TRY>


  Query Match             8.9%;  Score 77;  DB 2;  Length 603;
  Best Local Similarity   21.2%;  Pred. No. 6.1;
  Matches   39;  Conservative   18;  Mismatches   59;  Indels   68;  Gaps    8;

Qy      12 FHHTLQLGPREQARNASCPAGGRPADRRFRPPT-NLRSVSPWAYRISYDPARYPRY---- 66
           :||              |   |||  | :   | |      ||| :  :| :   :
Db      57 YHH-------------CIHKGRPGPRPWCATTPNFDQDQQWAYCL--EPKKVKDHCSKH 100

Qy      67 ------------LPEAYCLCRGCLTG---------------LFGEEDVRFRSAPV----- 94
                       |   :|||   |||               | | :: ||: |
Db     101 NPCQRGGICVNTLSSPHCLCPDHLTGKHCQREKCFEPQLHRFFHENEIWFRTGPAGVAKC 160

Qy      95 ----------YMPTVVLRRTPACAGGRSVYTEAY--VTIPVGCTCVPEPEKDADSINSSI 142
                    | :   :  |   ||| :  | :     |:| |    |  | |:  |
Db     161 HCKGPDAHCKQMHSQECQTNPCLNGGRCLEVEGHHLCDCPMGYT---GPFCDLDTTASCY 217

Qy     143 DKQG 146
           : :|
Db     218 EGRG 221


RESULT   8
B81719
conserved hypothetical protein TC0284 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: B81719
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255
A;Accession: B81719
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-242 <TET>
A;Cross-references: GB:AE002296; GB:AE002160; NID:g7190325; PIDN:AAF39152.1; PID:g719
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0284


  Query Match             8.8%;  Score 76;  DB 2;  Length 242;
  Best Local Similarity   26.2%;  Pred. No. 2.9;
  Matches   21;  Conservative    6;  Mismatches   21;  Indels   32;  Gaps    3;

Qy      68 PEAYCLCRGCLTGLF---------GEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYT--- 115
           || :| |  |  |             |||: |||  :                |  :||
Db     161 PEPHCNCLHCQIGRIAVEEEDIEVSEEDLTFRSWDI------------SQYGEKMYTVTD 208

Qy     116 --------EAYVTIPVGCTC 127
                     |:  |:||||
Db     209 PLNPEEQFNVYLGTPIGCTC 228


RESULT   9
T44254
thiamin biosynthesis protein thiC [imported] - Rhizobium etli plasmid b
C;Species: Rhizobium etli
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T44254
R;Miranda-Rios, J.; Morera, C.; Taboada, H.; Davalos, A.; Encarnacion, S.; Mora, J.;
J. Bacteriol. 179, 6887-6893, 1997
A;Title: Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiot
A;Reference number: Z22737; MUID:98037482
A;Accession: T44254
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-610 <MIR>
A;Cross-references: EMBL:AF004408; NID:g2627325; PIDN:AAC45972.1; PID:g2627326
A;Experimental source: strain CE3
C;Genetics:
A;Gene: thiC
A;Genome: plasmid b
C;Function:
A;Description: involved in the biosynthesis of the pyrimidine moiety of thiamin
C;Superfamily: thiamin biosynthesis protein thiC


  Query Match             8.7%;  Score 75.5;  DB 2;  Length 610;
  Best Local Similarity   29.9%;  Pred. No. 8.8;
  Matches   38;  Conservative   14;  Mismatches   54;  Indels   21;  Gaps    7;

Qy      30 PAGGRPADRRFRPPTNLRSVSPW---AYRISYDPARYPRYLPEAYCLCRGCLTGLFGE-- 84
           |  | |      | |   |  |:   |: || | |  || | |::   || :    |
Db      44 PTSGEP------PVTVYDSSGPYTDPAHVISID-AGLPR-LRESWIKARGDVESYDGRIV 95

Qy      85 --EDVRFRSAPVYMPTVVLRRTPACA-GGRSVYTEAYVTIPVGCTCVPEPEKDADSINSS 141
```

```
Db     96 KPEDNGFATGERLTPEFPVRNTPLKAKAGRAVTQLAYAR-----AGIVTPEMEFIAIREN 150

Qy    142 IDKQGAK 148
Db    151 LGRQAAK 157
```

RESULT 10
D72281
hypothetical protein TM1214 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: D72281
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316
A;Accession: D72281
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-178 <ARN>
A;Cross-references: GB:AE001778; GB:AE000512; NID:g4981757; PIDN:AAD36289.1; PID:g498176
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1214
C;Superfamily: psbG protein

Query Match 8.7%; Score 75; DB 2; Length 178;
Best Local Similarity 23.1%; Pred. No. 2.6;
Matches 34; Conservative 12; Mismatches 35; Indels 66; Gaps 7;

```
Qy     45 NLRSVSPWAYRISYDPARYPRYLPEAYCLCRGC--------LTGLFGEEDVRFRSAPV-- 94
Db     13 NLRSRSIWMLH----------------YCTGCGAVELPPSMTSRFDME--RFGIAPMAT 53

Qy     95 -----------YMPTVVLRRT----------------PACAGGRSVYTEAYVT------- 120
Db     54 PRQADILLITGYLNTKTLRRVIYTYEQMPDPKYVVGFGSCTINGGIYFDSYATVNRLDYY 113

Qy    121 IPVG---CTCVPEPEKDADSINSSIDK 144
Db    114 IPVDVYIAGCMPRPEAILEAFNYLMEK 140
```

RESULT 11
A71568
hypothetical protein CT016 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C;Accession: A71568
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809
A;Accession: A71568
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-242 <ARN>
A;Cross-references: GB:AE001276; GB:AE001273; NID:g3328399; PIDN:AAC67606.1; PID:g332840
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: CT016

Query Match 8.7%; Score 75; DB 2; Length 242;
Best Local Similarity 25.0%; Pred. No. 3.6;
Matches 20; Conservative 7; Mismatches 21; Indels 32; Gaps 3;

```
Qy     68 PEAYCLCRGCLTGL---------FGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYT--- 115
Db    161 PEPHCNCLHCQIGRATVEEEDAGVSDEDLTFRSWDI-----------SQSGEKMYTVTD 208

Qy    116 --------EAYVTIPVGCTC 127
Db    209 PLNPEEQFNVYLGTPIGCTC 228
```

RESULT 12
S22395
fetuin precursor - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jul-2000
C;Accession: S22395; S22142
R;Brown, W.M.; Dziegielewska, K.M.; Saunders, N.R.; Christie, D.L.; Nawratil, P.; Mue
Eur. J. Biochem. 205, 321-331, 1992
A;Title: The nucleotide and deduced amino acid structures of sheep and pig fetuin. Co
A;Reference number: S22394; MUID:92209519
A;Accession: S22395
A;Molecule type: mRNA
A;Residues: 1-362 <BRW>
A;Cross-references: EMBL:X56021; NID:g2104; PIDN:CAA39498.1; PID:g3980229
C;Superfamily: alpha-2-HS-glycoprotein; cystatin homology
C;Keywords: calcium binding; EF hand; glycoprotein
F;1-15/Domain: signal sequence (fragment) #status predicted <SIG>
F;16-362/Product: fetuin #status predicted <MAT>
F;20-134/Domain: cystatin homology <CY1>
F;143-249/Domain: cystatin homology <CY2>
F;96,153,173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.7%; Score 75; DB 2; Length 362;
Best Local Similarity 27.6%; Pred. No. 5.6;
Matches 27; Conservative 11; Mismatches 42; Indels 18; Gaps 3;

```
Qy     57 SYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTE 116
Db    210 AYSPTKCNLLVEKQYGFCKGTVTAKVNEEDVAVTCTVFQTQPVVLQPQPAGA-------- 261

Qy    117 AYVTIPVGCTCVPEPEKDADSINSSI-DKQGAKLLLGP 153
Db    262 -----DAGAT----PVVDAAATASPLADVPAASLVVGP 290
```

RESULT 13
I68674
cyclin-dependent kinase - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C;Accession: I68674
R;Mousses, S.; Ozcelik, H.; Lee, P.D.; Malkin, D.; Bull, S.B.; Andrulis, I.L.
Hum. Mol. Genet. 4, 1089-1092, 1995
A;Title: Two variants of the CIP1/WAF1 gene occur together and are associated with hu
A;Reference number: I54380; MUID:95384154
A;Accession: I68674
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-181 <RES>
A;Cross-references: GB:L47233; NID:g986878; PIDN:AAB59560.1; PID:g986879
C;Genetics:
A;Gene: CIP1/WAF1

Query Match 8.6%; Score 74.5; DB 2; Length 181;
Best Local Similarity 24.6%; Pred. No. 3;
Matches 43; Conservative 10; Mismatches 73; Indels 49; Gaps 6;

```
Qy      3 ARARAVLSAFHHTLQLGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPAR 62
Db      4 ARIRRGTEALRGGAMSEPAGDVRQNPC---GSKACRRLFGPVDSEQLS------------ 48

Qy     63 YPRYLPEAYCLCRGC----------------LTGLFGEEDVRFRSAP-VYMPTVVLRRT 104
```

```
Db     49 -----RDCDALMAGCIQEARERWNFDFVTETPLEGDFAWERVRGLGLPKLYLPTGPRRGR 103

Qy    105 PACAGGRSVYT----------EAYVTIPVGCTCVPEPEKDAD-SINSSIDKQGAK 148
Db    104 DELGGGRRPGTSPALLQGTAEEDHVDLSLSCTLVPRSGEQAEGSPGGPGDSQGRK 158


RESULT 14
T38292
hypothetical protein SPAC23E2.02 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T38292
R;Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z21784
A;Accession: T38292
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1273 <SKE>
A;Cross-references: EMBL:Z68887; PIDN:CAA93114.1; GSPDB:GN00066; SPDB:SPAC23E2.02
A;Experimental source: strain 972h-; cosmid c23E2
C;Genetics:
A;Gene: SPDB:SPAC23E2.02
A;Map position: 1
A;Introns: 8/2; 862/3


  Query Match             8.6%;  Score 74.5;  DB 2;  Length 1273;
  Best Local Similarity   25.6%;  Pred. No. 24;
  Matches   22;  Conservative   15;  Mismatches   48;  Indels    1;  Gaps    1;

Qy     73 LCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVGCTCVPEPE 132
Db    520 ICARQLTGLFSQYSSSFLSKNELPPKVIILEAKERTGGR-IYSRALPVSHTSATQINHHT 578

Qy    133 KDADSINSSIDKQGAKLLLGPNDAPA 158
Db    579 SNSNSISSNSTSLNPKDVTDPSHIPS 604


RESULT 15
RNFF2L
DNA-directed RNA polymerase (EC 2.7.7.6) II 215K chain [validated] - fruit fly (Drosophi
C;Species: Drosophila melanogaster
C;Date: 04-Dec-1986 #sequence_revision 31-Mar-1993 #text_change 28-Jul-2000
C;Accession: S04457; A00693; B27677; S60151
R;Jokerst, R.S.; Weeks, J.R.; Zehring, W.A.; Greenleaf, A.L.
Mol. Gen. Genet. 215, 266-275, 1989
A;Title: Analysis of the gene encoding the largest subunit of RNA polymerase II in Drosc
A;Reference number: S04457; MUID:89218930
A;Accession: S04457
A;Molecule type: DNA
A;Residues: 1-1896 <JOK>
A;Cross-references: EMBL:M27431; NID:g158331; PIDN:AAA28868.1; PID:g158332
R;Biggs, J.; Searles, L.L.; Greenleaf, A.L.
Cell 42, 611-621, 1985
A;Title: Structure of the eukaryotic transcription apparatus: features of the gene for t
A;Reference number: A00693; MUID:85282618
A;Accession: A00693
A;Molecule type: DNA
A;Residues: 1-318,'GYAKV',325-449,'G',451-454,'RCTT',459-462,'VTGESVASS' <BIG>
A;Cross-references: EMBL:M11798
A;Note: this sequence has been revised in reference S04457
A;Note: the authors translated the codon AAC for residue 451 as Thr
R;Allison, L.A.; Wong, J.K.C.; Fitzpatrick, V.D.; Moyle, M.; Ingles, C.J.
Mol. Cell. Biol. 8, 321-329, 1988
A;Title: The C-terminal domain of the largest subunit of RNA polymerase II of Saccharomy
A;Reference number: A93104; MUID:88094402
A;Accession: B27677
A;Molecule type: DNA
A;Residues: 1441-1484;'I',1527-1889 <ALL>
A;Cross-references: EMBL:M19537; NID:g158147; PIDN:AAA28827.1; PID:g158148
R;Petersen, G.; Song, D.; Huegle-Doerr, B.; Oldenburg, I.; Bautz, E.K.F.
Mol. Gen. Genet. 249, 425-431, 1995
A;Title: Mapping of linear epitopes recognized by monoclonal antibodies with gene-fra
A;Reference number: S60151; MUID:96133682
A;Accession: S60151
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 778-827 <PET>
C;Genetics:
A;Gene: RPO21; RpII215
A;Cross-references: FlyBase:FBgn0003277
A;Map position: X 10C, X 35.7
A;Introns: 27/3; 775/3; 1526/1
C;Function:
A;Description: EC 2.7.7.6 [validated; MUID:88094402]; essential for proper initiation
C;Superfamily: human DNA-directed RNA polymerase II largest chain
C;Keywords: DNA binding; nucleotidyltransferase; tandem repeat; transcription; zinc f
F;67-83/Region: zinc finger CCHH motif
F;1581-1883/Region: 7-residue repeats
F;349/Binding site: ATP/GTP (Lys) #status predicted


  Query Match             8.6%;  Score 74.5;  DB 1;  Length 1896;
  Best Local Similarity   24.8%;  Pred. No. 37;
  Matches   37;  Conservative   15;  Mismatches   50;  Indels   47;  Gaps    9;

Qy     20 PREQARNASCPAGGRPADRRFRPPT-NLRSVSPWAYRISYDPARY----PRYLPEAYCLC 74
Db   1583 PTSPNYTASSPGGASP---NYSPSSPNYSPTSP-----LYASPRYASTTPNFNPQS---- 1630

Qy     75 RGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAG-GRSVYT--EAYVTIPVGCTCVPEP 131
Db   1631 ----TGYSPSSSGYSPTSPVYSPTVQFQSSPSFAGSGSNIYSPGNAY----------SP 1676

Qy    132 EKDADSINSSIDKQGAKLLLGPNDAPAGP 160
Db   1676 SSSNYSPNS----------PSYSPTSP 1692


Search completed: February 16, 2001, 12:50:42
Job time: 148 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:52:57 ; Search time 14.73 Seconds
(without alignments)
350.784 Million cell updates/sec

Title: US-09-320-713-4
Perfect score: 864
Sequence: 1 NSARARAVLSAFHHTLQLGP.....SIDKQGAKLLLGPNDAPAGP 160

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	135.5	15.7	155	1	IL17_HUMAN	Q16552 homo sapien
2	113.5	13.1	158	1	IL17_MOUSE	Q62386 mus musculu
3	108.5	12.6	151	1	VG13_HSVSA	P24916 herpesvirus
4	108	12.5	150	1	IL17_RAT	Q61453 rattus norv
5	78.5	9.1	1075	1	NFC3_MOUSE	P97305 mus musculu
6	77	8.9	603	1	FA12_CAVPO	Q04962 cavia porce
7	75.5	8.7	610	1	THIC_RHIET	O34291 rhizobium e
8	75	8.7	362	1	A2HS_PIG	P29700 sus scrofa
9	74.5	8.6	1273	1	YAR2_SCHPO	Q10135 schizosacch
10	74.5	8.6	1896	1	RPB1_DROME	P04052 drosophila
11	73.5	8.5	1033	1	CR2_HUMAN	P20023 homo sapien
12	73	8.4	3358	1	PGCV_MOUSE	Q62059 mus musculu
13	71.5	8.3	164	1	CDN1_HUMAN	P38936 homo sapien
14	70.5	8.2	265	1	YM40_MYCTU	Q10522 mycobacteri
15	70.5	8.2	479	1	GLYC_NEUCR	P34898 neurospora
16	69.5	8.0	324	1	TTP_BOVIN	P53781 bos taurus
17	68.5	7.9	342	1	HXCA_MOUSE	P31257 mus musculu
18	68	7.9	2205	1	POLN_RUBVT	P13889 rubella vir
19	67.5	7.8	298	1	MYOD_CHICK	P16075 gallus gall
20	67.5	7.8	384	1	DUS9_HUMAN	Q99956 homo sapien
21	67.5	7.8	591	1	UL49_EBV	P14347 epstein-bar
22	67	7.8	277	1	HXDB_NOTVI	P31263 notophthalm
23	67	7.8	2273	1	HFA1_YEAST	P32874 saccharomyc
24	67	7.8	4393	1	PGBM_HUMAN	P98160 homo sapien
25	66.5	7.7	2339	1	CCAB_RABIT	Q05152 oryctolagus
26	66	7.6	489	1	VHS_HSV11	P10225 herpes simp
27	66	7.6	676	1	DNLJ_THETH	P26996 thermus aqu
28	65.5	7.6	170	1	PLGF_HUMAN	P49763 homo sapien
29	65.5	7.6	369	1	CHEB_RHOSH	O33558 rhodobacter
30	65.5	7.6	404	1	VE2_HPV60	Q80944 human papil
31	65.5	7.6	500	1	STCL_EMENI	Q00707 emericella
32	65	7.5	345	1	YJJN_ECOLI	P39400 escherichia
33	65	7.5	376	1	METK_AQUAE	O67222 aquifex aeo
34	65	7.5	802	1	ACSB_ACEXY	P37716 acetobacter
35	65	7.5	1106	1	GLI1_HUMAN	P08151 homo sapien
36	64.5	7.5	285	1	HXA4_MOUSE	P06798 mus musculu
37	64.5	7.5	374	1	RLIG_BPT4	P00971 bacteriopha
38	64	7.4	286	1	PYR1_SYNEL	P50034 synechococc
39	64	7.4	453	1	CARA_NEUCR	P22572 neurospora
40	64	7.4	497	1	CPDH_MACFA	Q29488 macaca fasc
41	64	7.4	3396	1	PGCV_HUMAN	P13611 homo sapien
42	63.5	7.3	258	1	UL49_HSVBC	P30022 bovine herp
43	63.5	7.3	419	1	IRF3_MOUSE	P70671 mus musculu
44	63.5	7.3	429	1	ELK1_MOUSE	P41969 mus musculu
45	63.5	7.3	476	1	C6B5_PAPGL	Q95036 papilio gla

ALIGNMENTS

```
RESULT  1
IL17_HUMAN
ID   IL17_HUMAN     STANDARD;      PRT;   155 AA.
AC   Q16552;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED
DE   ANTIGEN 8) (CTLA-8).
GN   IL17 OR CTLA8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96281911; PubMed=8676080;
RA   Fossiez F., Djossou O., Chomarat P., Flores-Romo L., Ait-Yahia S.,
RA   Maat C., Pin J.-J., Garrone P., Garcia E., Saeland S., Blanchard D.,
RA   Gaillard C., Das Mahapatra B., Rouvier E., Golstein P., Banchereau J.,
RA   Lebecque S.;
RT   "T cell interleukin-17 induces stromal cells to produce
RT   proinflammatory and hematopoietic cytokines.";
RL   J. Exp. Med. 183:2593-2603(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T-CELL;
RX   MEDLINE=96094436; PubMed=7499828;
RA   Yao Z., Painter S.L., Fanslow W.C., Ulrich D., Macduff B.M.,
RA   Spriggs M.K., Armitage R.J.;
RT   "Human IL-17: a novel cytokine derived from T cells.";
RL   J. Immunol. 155:5483-5486(1995).
CC   -!- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND
CC       HEMATOPOIETIC CYTOKINES. ENHANCES THE SURFACE EXPRESSION OF THE
CC       INTRACELLULAR ADHESION MOLECULE-1 (ICAM-1) IN FIBROBLASTS.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO ACTIVATED MEMORY T-CELLS.
CC   -!- PTM: FOUND BOTH IN GLYCOSYLATED AND NONGLYCOSYLATED FORMS.
CC   -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO
CC       HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z58820; CAA91233.1; -.
DR   EMBL; U32659; AAC50341.1; -.
DR   MIM; 603149; -.
KW   Cytokine; Glycoprotein; T-cell; Antigen; Signal.
FT   SIGNAL       1     23       POTENTIAL.
FT   CHAIN       24    155       INTERLEUKIN-17.
FT   CARBOHYD    68     68       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
SQ   SEQUENCE   155 AA;  17504 MW;  2BCAE9CB2F4886D1 CRC64;

  Query Match             15.7%;  Score 135.5;  DB 1;  Length 155;
  Best Local Similarity   32.3%;  Pred. No. 4e-07;
  Matches   41;  Conservative    9;  Mismatches   46;  Indels   31;  Gaps    7;

Qy      25 RNASCPAGGRPADRRFRPPT---NL------------------RSVSPWAYRISYDPAR 62
           ||  ||      |: | | |   ||                  || |||    : || |
Db      29 RNPGCP---NSEDKNF-PRTVMVNLNIHNRNTNTNPKRSSDYYNRSTSPWNLHRNEDPER 84

Qy      63 YPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA-CAGGRSVYTEAYVTI 121
           ||  : || |   ||:    | |    | |:    :|||| |  |      :    : :
Db      85 YPSVIWEAKCRHLGCINA-DGNVDYHMNSVPIQQEILVLRREPPHCPNS---FRLEKILV 140

Qy     122 PVGCTCV 128
           |||||||
Db     141 SVGCTCV 147


RESULT   2
IL17_MOUSE
ID   IL17_MOUSE     STANDARD;      PRT;   158 AA.
AC   Q62386; Q60971;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED
DE   ANTIGEN 8) (CTLA-8).
GN   IL17 OR CTLA8.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=THYMOCYTES;
RX   MEDLINE=97031826; PubMed=8877732;
RA   Kennedy J., Rossi D.L., Zurawski S.M., Vega F. Jr., Kastelein R.A.,
RA   Wagner J.L., Hannum C.H., Zlotnik A.;
RT   "Mouse IL-17: a cytokine preferentially expressed by alpha beta TCR +
RT   CD4-CD8-T cells.";
RL   J. Interferon Cytokine Res. 16:611-617(1996).
RN   [2]
RP   SEQUENCE OF 12-158 FROM N.A.
RC   STRAIN=129/SV; TISSUE=T-CELL;
RX   MEDLINE=96194901; PubMed=8654948;
RA   Yao Z., Timour M., Painter S., Fanslow W., Spriggs M.K.;
RT   "Complete nucleotide sequence of the mouse CTLA8 gene.";
RL   Gene 168:223-225(1996).
CC   -!- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND
CC       HEMATOPOIETIC CYTOKINES (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO A SUBSET OF ACTIVATED T-CELLS.
CC   -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO
CC       HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U43088; AAB05222.1; -.
DR   EMBL; U35108; AAA93253.1; -.
DR   MGD; MGI:107364; IL17.
KW   Cytokine; Glycoprotein; T-cell; Antigen; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    158       INTERLEUKIN-17.
FT   CARBOHYD     71     71       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   158 AA;  17490 MW;  3505C143435F4653 CRC64;

  Query Match             13.1%;  Score 113.5;  DB 1;  Length 158;
  Best Local Similarity   31.4%;  Pred. No. 7.7e-05;
  Matches   37;  Conservative   16;  Mismatches   46;  Indels   19;  Gaps    6;

Qy      32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFR 90
           | : : |  ||    | || |||    : || |||  : || |  : |:    |: |
Db      58 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMN 114

Qy      91 SAPVYMPTVVLRRTP-ACAGGRSVYTEAYVTIPVGCTCVPEPEKDADSINSSIDKQGA 147
           |  :    :||:| | :|      :    : : ||||||           :|| :| |
Db     115 SVLIQQEILVLKREPESCP---FTFRVEKMLVGVGCTCV----------ASIVRQAA 158


RESULT   3
VG13_HSVSA
ID   VG13_HSVSA     STANDARD;      PRT;   151 AA.
AC   P24916;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   IMMEDIATE EARLY GENE 13 PROTEIN PRECURSOR.
GN   13 OR KCLF2.
OS   Herpesvirus saimiri (strain 11).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90163221; PubMed=2154888;
RA   Albrecht J.-C., Fleckenstein B.;
RT   "Structural organization of the conserved gene block of Herpesvirus
RT   saimiri coding for DNA polymerase, glycoprotein B, and major DNA
RT   binding protein.";
RL   Virology 174:533-542(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92333688; PubMed=1321287;
RA   Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA   Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA   Honess R.W.;
RT   "Primary structure of the herpesvirus saimiri genome.";
RL   J. Virol. 66:5047-5058(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91021021; PubMed=1699352;
RA   Nicholas J., Smith E.P., Coles L., Honess R.;
RT   "Gene expression in cells infected with gammaherpesvirus saimiri:
RT   properties of transcripts from two immediate-early genes.";
RL   Virology 179:189-200(1990).
CC   -!- SIMILARITY: STRONG, TO MAMMALIAN INTERLEUKIN-17 (CTLA-8).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X64346; CAA45636.1; -.
DR   EMBL; M31122; AAA46169.1; -.
DR   EMBL; M60286; AAA46156.1; -.
DR   PIR; D36807; D36807.
DR   PIR; B45351; B45351.
KW   Cytokine; Early protein; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    151       IMMEDIATE EARLY GENE 13 PROTEIN.
FT   CARBOHYD     36     36       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     53     53       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   151 AA;  17180 MW;  53BEDDE4206C6432 CRC64;
```

```
  Query Match             12.6%;  Score 108.5;  DB 1;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 0.00024;
  Matches   27;  Conservative   11;  Mismatches   37;  Indels   11;  Gaps    3;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLHRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
               |      ::    : : ||||||
Db     121 ---GHQPCPNSFRLEKMLVTVGCTCV 143


RESULT   4
IL17_RAT
ID   IL17_RAT       STANDARD;      PRT;   150 AA.
AC   Q61453;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED
DE   ANTIGEN 8) (CTLA-8).
GN   IL17 OR CTLA8.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93294300; PubMed=8390535;
RA   Rouvier E., Luciani M.F., Mattei M.-G., Denizot F., Golstein P.;
RT   "CTLA-8, cloned from an activated T cell, bearing AU-rich messenger
RT   RNA instability sequences, and homologous to a herpesvirus saimiri
RT   gene.";
RL   J. Immunol. 150:5445-5456(1993).
RN   [2]
RP   ORGANISM IDENTIFICATION.
RX   MEDLINE=96194901; PubMed=8654948;
RA   Yao Z., Timour M., Painter S., Fanslow W., Spriggs M.K.;
RT   "Complete nucleotide sequence of the mouse CTLA8 gene.";
RL   Gene 168:223-225(1996).
RN   [3]
RP   SEQUENCE FROM N.A., AND ORGANISM IDENTIFICATION.
RX   MEDLINE=97031826; PubMed=8877732;
RA   Kennedy J., Rossi D.L., Zurawski S.M., Vega F. Jr., Kastelein R.A.,
RA   Wagner J.L., Hannum C.H., Zlotnik A.;
RT   "Mouse IL-17: a cytokine preferentially expressed by alpha beta TCR +
RT   CD4-CD8-T cells.";
RL   J. Interferon Cytokine Res. 16:611-617(1996).
CC   -!- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND
CC       HEMATOPOIETIC CYTOKINES (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO
CC       HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.
CC   -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE FROM MOUSE BUT, ON
CC       THE BASIS OF SUBSEQUENT WORK (REF.2 AND REF.3), HAS BEEN SHOWN TO
CC       BE OF RAT ORIGIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L13839; AAA37490.1; -.
KW   Cytokine; Glycoprotein; T-cell; Antigen; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    150       INTERLEUKIN-17.
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     46     46       I -> L (IN REF. 3).
SQ   SEQUENCE   150 AA;  16876 MW;  EF13F33EDF9D689F CRC64;


  Query Match             12.5%;  Score 108;  DB 1;  Length 150;
  Best Local Similarity   34.4%;  Pred. No. 0.00027;
  Matches   31;  Conservative   11;  Mismatches   42;  Indels    6;  Gaps    4;

Qy      41 RPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTV 99
           ||   | || |||    : || |||  : || |  : |:    |: |    |  :    :
Db      57 RPSDYLNRSTSPWTLSRNEDPDRYPSVIWEAQCRHQRCVNA-EGKLDHHMNSVLIQQEIL 115

Qy     100 VLRRTP-ACAGGRSVYTEAYVTIPVGCTCV 128
           ||:| |  |      :    : : ||||||
Db     116 VLKREPEKCP---FTFRVEKMLVGVGCTCV 142


RESULT   5
NFC3_MOUSE
ID   NFC3_MOUSE     STANDARD;      PRT;  1075 AA.
AC   P97305; Q60896;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   NUCLEAR FACTOR OF ACTIVATED T-CELLS, CYTOPLASMIC 3 (T CELL
DE   TRANSCRIPTION FACTOR NFAT4) (NF-ATC3) (NF-AT4) (NFATX).
GN   NFATC3 OR NFAT4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM X1).
RC   TISSUE=THYMUS;
RX   MEDLINE=95378239; PubMed=7650004;
RA   Ho S.N., Thomas D.J., Timmerman L.A., Li X., Francke U.,
RA   Crabtree G.R.;
RT   "NFATc3, a lymphoid-specific NFATc family member that is
RT   calcium-regulated and exhibits distinct DNA binding specificity.";
RL   J. Biol. Chem. 270:19898-19907(1995).
RN   [2]
RP   SEQUENCE OF 12-1075 FROM N.A. (ISOFORMS X1; X2 AND DELTA-X).
RC   TISSUE=THYMIC LYMPHOMA;
RX   MEDLINE=97170074; PubMed=9017603;
RA   Liu J., Koyano-Nakagawa N., Amasaki Y., Saito-Ohara F., Ikeuchi T.,
RA   Imai S.-I., Takano T., Arai N., Yokota T., Arai K.-I.;
RT   "Calcineurin-dependent nuclear translocation of a murine transcription
RT   factor NFATx: molecular cloning and functional characterization.";
RL   Mol. Biol. Cell 8:157-170(1997).
RN   [3]
RP   REVIEW.
RX   MEDLINE=99189746; PubMed=10089876;
RA   Crabtree G.R.;
RT   "Generic signals and specific outcomes: signaling through Ca2+,
RT   calcineurin, and NF-AT.";
RL   Cell 96:611-614(1999).
CC   -!- FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE
CC       GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2 (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: MEMBER OF THE MULTICOMPONENT NFATC TRANSCRIPTION COMPLEX
CC       THAT CONSISTS OF AT LEAST TWO COMPONENTS, A PRE-EXISTING
CC       CYTOPLASMIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT
CC       NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC3 OR MEMBERS OF THE
CC       ACTIVATING PROTEIN-1 FAMILY, MAF, GATA4 AND CBP/300 CAN ALSO BIND
CC       THE COMPLEX. NFATC PROTEINS BIND TO DNA AS MONOMERS.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC FOR THE PHOSPHORYLATED FORM AND
CC       NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-
CC       MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT
CC       TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED
CC       AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF
CC       NFATC PLAY A KEY ROLE IN THE GENE TRANSCRIPTION.
CC   -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; X1 (SHOWN HERE), X2 AND DELTA-X;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THYMUS. WEAKLY EXPRESSED IN
CC       MUSCLE, SPLEEN AND KIDNEY. ALSO EXPRESSED IN LYMPH NODE.
```

```
CC   -!- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND
CC       COOPERATIVE INTERACTIONS WITH AP1 FACTORS (BY SIMILARITY).
CC   -!- PTM: PHOSPHORYLATED BY NFATC-KINASE; DEPHOSPHORYLATED BY
CC       CALCINEURIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D85612; BAA12833.1; -.
DR   EMBL; U28807; AAA93249.1; -.
DR   MGD; MGI:103296; NFATC3.
DR   INTERPRO; IPR000451; -.
DR   PROSITE; PS01204; REL_1; FALSE_NEG.
DR   PROSITE; PS50254; REL_2; 1.
KW   Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW   Alternative splicing; Repeat; Phosphorylation.
FT   DOMAIN       24     29       POLY-PRO.
FT   DOMAIN      109    114       CALCINEURIN-BINDING.
FT   DOMAIN      207    308       3 X SP REPEATS.
FT   REPEAT      207    223       1.
FT   REPEAT      236    252       2.
FT   REPEAT      292    308       3.
FT   DOMAIN      273    275       NUCLEAR LOCALIZATION SIGNAL.
FT   DOMAIN      444    451       DNA-BINDING.
FT   DOMAIN      686    688       NUCLEAR LOCALIZATION SIGNAL.
FT   DOMAIN     1031   1040       NUCLEAR EXPORT SIGNAL.
FT   VARSPLIC    468    497       MISSING (IN ISOFORM DELTA-X).
FT   VARSPLIC   1035   1075       VNEIIGRDMSQISVSQATEVMRDTPLPGPASPDLMTSHSAH
FT                                -> DQLISDLEHQPSGSTEKWSNHSEFSCPVPFWRI (IN
FT                                ISOFORM X2).
FT   CONFLICT     12     34       LDFKLVFGEDGAPAPPPPGSRPA -> PRRVLFSVSAQLPS
FT                                RTRPGPSDL (IN REF. 2).
FT   CONFLICT     57     61       TTPLC -> NSSLG (IN REF. 2).
FT   CONFLICT     68     68       P -> LQ (IN REF. 2).
FT   CONFLICT     82     83       SH -> GY (IN REF. 2).
FT   CONFLICT     90     93       CEIP -> GDIS (IN REF. 2).
FT   CONFLICT    113    113       F -> I (IN REF. 2).
FT   CONFLICT    121    124       QQEL -> HQGT (IN REF. 2).
FT   CONFLICT    132    132       Q -> H (IN REF. 2).
FT   CONFLICT    140    140       F -> Y (IN REF. 2).
FT   CONFLICT    641    641       G -> R (IN REF. 2).
FT   CONFLICT    646    646       V -> A (IN REF. 2).
FT   CONFLICT    707    734       REDTDLPSVPSLPVPHSAQAQRPSSETG -> KEKTQICLQ
FT                                FHLCLCLILPSPEALLRDR (IN REF. 2).
FT   CONFLICT    746    746       G -> S (IN REF. 2).
FT   CONFLICT   1075   1075       H -> Q (IN REF. 2).
SQ   SEQUENCE   1075 AA;  115450 MW;  4ED38C9AA6F452BB CRC64;


  Query Match             9.1%;  Score 78.5;  DB 1;  Length 1075;
  Best Local Similarity   22.0%;  Pred. No. 2.8;
  Matches   37;  Conservative   22;  Mismatches   76;  Indels   33;  Gaps    5;

Qy       1 NSARARAVLSAFHHTLQLGPREQARNASCPA-----------------GGRPADRRFRP 42
           |  | ::    | :|  |  :|| :   |:                   | | ||
Db     684 NGKRKKSQSQRFTYTPVLMKQEQREDTDLPSVPSLPVPHSAQAQRPSSETGHPHDRAMSA 743

Qy      43 PTNLR-SVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVL 101
           |  |   | |     :|       :||:  |   |      |:|     |: :: |  |
Db     744 PGGLLCQVQP-----AYTSMVASTHLPQLQCRDEGA-----GKEQHIATSSVMHQPFQVT 793

Qy     102 RRTPACAGGRSVYTEAYVTIPVGCTCVPEPEKDADSINSSIDKQGAKL 149
             :|  :  :|: |  |     | ||:|     :   :  : :| | |
Db     794 PTSPIGSSYQSIQTSMY----NGPTCLPVNPASSQEFDPVLFQQDAAL 837


RESULT   6
```

```
FA12_CAVPO
ID   FA12_CAVPO     STANDARD;      PRT;   603 AA.
AC   Q04962;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   COAGULATION FACTOR XII PRECURSOR (EC 3.4.21.38) (HAGEMAN FACTOR)
DE   (HAF) (FRAGMENT).
GN   F12.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC   TISSUE=LIVER;
RX   MEDLINE=93003367; PubMed=1390917;
RA   Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
RA   Kambara T., Okabe H.;
RT   "Primary structure of guinea-pig Hageman factor: sequence around the
RT   cleavage site differs from the human molecule.";
RL   Biochim. Biophys. Acta 1159:113-121(1992).
CC   -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC       THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC       GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC   -!- CATALYTIC ACTIVITY: CLEAVES SELECTIVELY ARG-|-ILE BONDS AND
CC       ACTIVATES COAGULATION FACTORS VII AND XI.
CC   -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC       COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC       FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC       TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC       XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE-I DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE-II DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 KRINGLE REGION.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X68615; CAA48600.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.211; -.
DR   INTERPRO; IPR000001; -.
DR   INTERPRO; IPR000083; -.
DR   INTERPRO; IPR000561; -.
DR   INTERPRO; IPR000562; -.
DR   INTERPRO; IPR001254; -.
DR   PFAM; PF00008; EGF; 2.
DR   PFAM; PF00039; fn1; 1.
DR   PFAM; PF00040; fn2; 1.
DR   PFAM; PF00051; kringle; 1.
DR   PFAM; PF00089; trypsin; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01253; FIBRONECTIN_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 1.
KW   Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW   Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     18
FT   CHAIN         19    358       ALPHA-FACTOR XIIA HEAVY CHAIN.
FT   CHAIN        359    603       ALPHA-FACTOR XIIA LIGHT CHAIN.
```

```
FT   DOMAIN       46     87       FIBRONECTIN TYPE-II.
FT   DOMAIN       93    130       EGF-LIKE 1.
FT   DOMAIN      132    172       FIBRONECTIN TYPE-I.
FT   DOMAIN      173    209       EGF-LIKE 2.
FT   DOMAIN      216    294       KRINGLE.
FT   DOMAIN      312    342       PRO-RICH.
FT   DOMAIN      359    603       CATALYTIC.
FT   ACT_SITE    398    398       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    447    447       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    551    551       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     97    109       BY SIMILARITY.
FT   DISULFID    103    118       BY SIMILARITY.
FT   DISULFID    120    129       BY SIMILARITY.
FT   DISULFID    134    162       BY SIMILARITY.
FT   DISULFID    160    169       BY SIMILARITY.
FT   DISULFID    177    188       BY SIMILARITY.
FT   DISULFID    182    197       BY SIMILARITY.
FT   DISULFID    199    208       BY SIMILARITY.
FT   DISULFID    216    294       BY SIMILARITY.
FT   DISULFID    237    276       BY SIMILARITY.
FT   DISULFID    265    289       BY SIMILARITY.
FT   DISULFID    345    472       BY SIMILARITY.
FT   DISULFID    383    399       BY SIMILARITY.
FT   DISULFID    391    461       BY SIMILARITY.
FT   DISULFID    422    425       BY SIMILARITY.
FT   DISULFID    488    557       BY SIMILARITY.
FT   DISULFID    520    536       BY SIMILARITY.
FT   DISULFID    547    578       BY SIMILARITY.
FT   CARBOHYD    248    248       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    270    270       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    419    419       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   603 AA;  66795 MW;  48DC6B946FB9ED59 CRC64;
```

```
  Query Match             8.9%;  Score 77;  DB 1;  Length 603;
  Best Local Similarity   21.2%;  Pred. No. 2.1;
  Matches   39;  Conservative   18;  Mismatches   59;  Indels   68;  Gaps    8;

Qy      12 FHHTLQLGPREQARNASCPAGGRPADRRFRPPT-NLRSVSPWAYRISYDPARYPRY---- 66
           :||              |   |||  | :   | |      |:| :  :| :   :
Db      57 YHH-------------CIHKGRPGPRPWCATTPNFDQDQQWAYCL--EPKKVKDHCSKH 100

Qy      67 ------------LPEAYCLCRGCLTG--------------LFGEEDVRFRSAPV----- 94
                      |   :|||   |||                | | :: ||: |
Db     101 NPCQRGGICVNTLSSPHCLCPDHLTGKHCQREKCFEPQLHRFFHENEIWFRTGPAGVAKC 160

Qy      95 ----------YMPTVVLRRTPACAGGRSVYTEAY--VTIPVGCTCVPEPEKDADSINSSI 142
                    | :   :  |   ||| :  | :      |:| |    |  | |:  |
Db     161 HCKGPDAHCKQMHSQECQTNPCLNGGRCLEVEGHHLCDCPMGYT---GPFCDLDTTASCY 217

Qy     143 DKQG 146
           : :|
Db     218 EGRG 221
```

```
RESULT   7
THIC_RHIET
ID   THIC_RHIET     STANDARD;      PRT;   610 AA.
AC   O34291;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   THIAMINE BIOSYNTHESIS PROTEIN THIC.
GN   THIC.
OS   Rhizobium etli.
OG   Plasmid pb.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CE3;
RX   MEDLINE=98037482; PubMed=9371431;
RA   Miranda-Rios J., Morera C., Taboada H., Davalos A., Encarnacion S.,
RA   Mora J., Soberon M.;
RT   "Expression of thiamin biosynthetic genes (thiCOGE) and production of
RT   symbiotic terminal oxidase cbb3 in Rhizobium etli.";
RL   J. Bacteriol. 179:6887-6893(1997).
CC   -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE HYDROMETHYLPYRIMIDINE
CC       (HMP) MOIETY OF THIAMINE (4-AMINO-2-METHYL-5-
CC       HYDROXYMETHYLPYRIMIDINE) (BY SIMILARITY).
CC   -!- PATHWAY: THIAMINE BIOSYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO THE THIC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF004408; AAC45972.1; -.
DR   INTERPRO; IPR002817; -.
DR   PFAM; PF01964; ThiC; 1.
KW   Thiamine biosynthesis; Plasmid.
SQ   SEQUENCE   610 AA;  67105 MW;  9CE7F560DB35ACDC CRC64;
```

```
  Query Match             8.7%;  Score 75.5;  DB 1;  Length 610;
  Best Local Similarity   29.9%;  Pred. No. 3;
  Matches   38;  Conservative   14;  Mismatches   54;  Indels   21;  Gaps    7;

Qy      30 PAGGRPADRRFRPPTNLRSVSPW---AYRISYDPARYPRYLPEAYCLCRGCLTGLFGE--- 84
           |  | |      | |   |  |:   |: || | |  || | |::   || :    |
Db      44 PTSGEP------PVTVYDSSGPYTDPAHVISID-AGLPR-LRESWIKARGDVESYDGRIV 95

Qy      85 --EDVRFRSAPVYMPTVVLRRTPACA-GGRSVYTEAYVTIPVGCTCVPEPEKDADSINSS 141
             ||  | :     |   :| ||  |  ||:|   ||         :  || :  :|  :
Db      96 KPEDNGFATGERLTPEFPVRNTPLKAKAGRAVTQLAYAR-----AGIVTPEMEFIAIREN 150

Qy     142 IDKQGAK 148
           : :| ||
Db     151 LGRQAAK 157
```

```
RESULT   8
A2HS_PIG
ID   A2HS_PIG       STANDARD;      PRT;   362 AA.
AC   P29700;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   ALPHA-2-HS-GLYCOPROTEIN PRECURSOR (FETUIN) (FRAGMENT).
GN   AHSG.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=LIVER;
RX   MEDLINE=92209519; PubMed=1372866;
RA   Brown W.M., Christie D.L., Saunders N.R., Nawratil P.,
RA   Dziegielewska K.D., Mueller-Esterl W.;
RT   "The nucleotide and deduced amino acid structures of sheep and pig
RT   fetuin. Common structural features of the mammalian fetuin family!";
RL   Eur. J. Biochem. 205:321-331(1992).
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE FETUIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 CYSTATIN-LIKE REPEATS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X56021; CAA39498.1; -.
DR   PIR; S22142; S22142.
DR   PIR; S22395; S22395.
DR   INTERPRO; IPR000010; -.
DR   INTERPRO; IPR001363; -.
DR   PFAM; PF00031; cystatin; 2.
DR   PROSITE; PS01254; FETUIN_1; 1.
DR   PROSITE; PS01255; FETUIN_2; 1.
KW   Glycoprotein; Signal; Repeat.
FT   NON_TER       1      1
FT   SIGNAL       <1     15
FT   CHAIN         16    362       ALPHA-2-HS-GLYCOPROTEIN.
FT   REPEAT        24    141       CYSTATIN-LIKE.
FT   REPEAT       142    258       CYSTATIN-LIKE.
FT   DISULFID      29    353       INTERCHAIN (BY SIMILARITY).
FT   DISULFID      86     97       BY SIMILARITY.
FT   DISULFID     111    129       BY SIMILARITY.
FT   DISULFID     143    146       BY SIMILARITY.
FT   DISULFID     205    216       BY SIMILARITY.
FT   DISULFID     227    244       BY SIMILARITY.
FT   CARBOHYD      96     96       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     153    153       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     173    173       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   362 AA;  38424 MW;  532648EE434B5686 CRC64;


  Query Match             8.7%;  Score 75;  DB 1;  Length 362;
  Best Local Similarity   27.6%;  Pred. No. 1.9;
  Matches   27;  Conservative   11;  Mismatches   42;  Indels   18;  Gaps    3;

Qy      57 SYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTE 116
           :| | :    : : |  |:| :|    ||||           |||:  || |
Db     210 AYSPTKCNLLVEKQYGFCKGTVTAKVNEEDVAVTCTVFQTQPVVLQPQPAGA-------- 261

Qy     117 AYVTIPVGCTCVPEPEKDADSINSSI-DKQGAKLLLGP 153
               | |    |  || :  | : |   | |::||
Db     262 -----DAGAT----PVVDAAATASPLADVPAASLVVGP 290


RESULT   9
YAR2_SCHPO
ID   YAR2_SCHPO     STANDARD;      PRT;  1273 AA.
AC   Q10135;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   HYPOTHETICAL 142.5 KDA PROTEIN C23E2.02 IN CHROMOSOME I.
GN   SPAC23E2.02.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC   Schizosaccharomycetaceae; Schizosaccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z68887; CAA93114.1; -.
DR   INTERPRO; IPR002937; -.
DR   PFAM; PF01593; Amino_oxidase; 1.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM   1028   1048       POTENTIAL.
SQ   SEQUENCE   1273 AA;  142489 MW;  3FEFF6433DF3C0B6 CRC64;


  Query Match             8.6%;  Score 74.5;  DB 1;  Length 1273;
  Best Local Similarity   25.6%;  Pred. No. 8.8;
  Matches   22;  Conservative   15;  Mismatches   48;  Indels    1;  Gaps    1;

Qy      73 LCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVGCTCVPEPE 132
           :|   ||||| :    | |     | |::       ||| :|: |        | :
Db     520 ICARQLTGLFSQYSSSFLSKNELPPKVIILEAKERTGGR-IYSRALPVSHTSATQINHHT 578

Qy     133 KDADSINSSIDKQGAKLLLGPNDAPA 158
           :::||:|:      | :  |:  |:
Db     579 SNSNSISSNSTSLNPKDVTDPSHIPS 604


RESULT  10
RPB1_DROME
ID   RPB1_DROME     STANDARD;      PRT;  1896 AA.
AC   P04052;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT (EC 2.7.7.6).
GN   RPII215.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89218930; PubMed=2496296;
RA   Jokerst R.S., Weeks J.R., Zehring W.A., Greenleaf A.L.;
RT   "Analysis of the gene encoding the largest subunit of RNA polymerase
RT   II in Drosophila.";
RL   Mol. Gen. Genet. 215:266-275(1989).
RN   [2]
RP   SEQUENCE OF 1-472 FROM N.A.
RX   MEDLINE=85282618; PubMed=2992806;
RA   Biggs J., Searles L.L., Greenleaf A.L.;
RT   "Structure of the eukaryotic transcription apparatus: features of the
RT   gene for the largest subunit of Drosophila RNA polymerase II.";
RL   Cell 42:611-621(1985).
RN   [3]
RP   SEQUENCE OF 1441-1889 FROM N.A.
RX   MEDLINE=88094402; PubMed=3122024;
RA   Allison L.A., Wong J.K.-C., Fitzpatrick V.D., Moyle M., Ingles C.J.;
RT   "The C-terminal domain of the largest subunit of RNA polymerase II of
RT   Saccharomyces cerevisiae, Drosophila melanogaster, and mammals: a
RT   conserved structure with an essential function.";
RL   Mol. Cell. Biol. 8:321-329(1988).
CC   -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC       OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC       SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
CC       RNA(N).
CC   -!- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO
CC       TWELVE DIFFERENT POLYPEPTIDES. THE 215 KDA POLYPEPTIDE IS THE
CC       LARGEST COMPONENT OF RNA POLYMERASE II.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC       THE PHOSPHORYLATION ACTIVATES POL2.
CC   -!- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
CC       FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC       PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC       III FOR 5S AND TRNA GENES.
CC   -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M27431; AAA28868.1; -.
DR   EMBL; M11798; AAA28863.1; -.
DR   EMBL; M19537; AAA28827.1; -.
DR   PIR; S04457; RNFF2L.
DR   FLYBASE; FBgn0003277; RpII215.
DR   INTERPRO; IPR000684; -.
DR   INTERPRO; IPR000722; -.
DR   INTERPRO; IPR002879; -.
DR   PFAM; PF01854; RNA_pol_A2; 1.
DR   PFAM; PF00623; RNA_pol_A; 1.
DR   PROSITE; PS00115; RNA_POL_II_REPEAT; 11.
KW   Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW   DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT   ZN_FING      67     83       C2H2-TYPE (POTENTIAL).
FT   DOMAIN     1581   1883       CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
FT   CONFLICT    319    324       RAMQKS -> GYAKV (IN REF. 2).
FT   CONFLICT    450    450       F -> G (IN REF. 2).
FT   CONFLICT    455    458       TLHK -> RCTT (IN REF. 2).
FT   CONFLICT    463    472       GHRVKVLPWS -> VTGESVASST (IN REF. 2).
SQ   SEQUENCE   1896 AA;  210045 MW;  801EA661FC0BAD5B CRC64;


  Query Match             8.6%;  Score 74.5;  DB 1;  Length 1896;
  Best Local Similarity   24.8%;  Pred. No. 14;
  Matches   37;  Conservative   15;  Mismatches   50;  Indels   47;  Gaps    9;

Qy      20 PREQARNASCPAGGRPADRRFRPPT-NLRSVSPWAYRISYDPARY----PRYLPEAYCLC 74
           |      || | |  |    : | : |    ||       |   ||    | : |::
Db    1583 PTSPNYTASSPGGASP---NYSPSSPNYSPTSP-----LYASPRYASTTPNFNPQS---- 1630

Qy      75 RGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAG-GRSVYT--EAYVTIPVGCTCVPEP 131
             ||              ::||| |||  : :|: || | ::|:   ||           |
Db    1631 ----TGYSPSSSGYSPTSPVYSPTVQFQSSPSFAGSGSNIYSPGNAY----------SP 1675

Qy     132 EKDADSINSSIDKQGAKLLLGPNDAPAGP 160
               | ||           |: :|  |
Db    1676 SSSNYSPNS-----------PSYSPTSP 1692


RESULT  11
CR2_HUMAN
ID   CR2_HUMAN      STANDARD;      PRT;  1033 AA.
AC   P20023;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   COMPLEMENT RECEPTOR TYPE 2 PRECURSOR (CR2) (COMPLEMENT C3D RECEPTOR)
DE   (EPSTEIN-BARR VIRUS RECEPTOR) (CD21 ANTIGEN).
GN   CR2 OR C3DR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89123277; PubMed=2563370;
RA   Fujisaku A., Harley J.B., Frank M.B., Gruner B.A., Frazier B.,
RA   Holers V.M.;
RT   "Genomic organization and polymorphisms of the human C3d/Epstein-Barr
RT   virus receptor.";
RL   J. Biol. Chem. 264:2118-2125(1989).
RN   [2]
RP   SEQUENCE OF 226-233; 256-267; 332-341; 667-677 AND 898-908.
RX   MEDLINE=86287311; PubMed=3016712;
RA   Weis J.J., Fearon D.T., Klickstein L.B., Wong W.W., Richards S.A.,
RA   de Bruyn Kops A., Smith J.A., Weis J.H.;
RT   "Identification of a partial cDNA clone for the C3d/Epstein-Barr
RT   virus receptor of human B lymphocytes: homology with the receptor for
RT   fragments C3b and C4b of the third and fourth components of
RT   complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5639-5643(1986).
CC   -!- FUNCTION: RECEPTOR FOR COMPLEMENT C3D AND FOR THE EPSTEIN-BARR
CC       VIRUS ON HUMAN B-CELLS. PARTICIPATES IN B LYMPHOCYTES ACTIVATION.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: MATURE B LYMPHOCYTES AND FOLLICULAR
CC       DENDRITIC CELLS OF THE SPLEEN.
CC   -!- SIMILARITY: TO MOUSE CR2. CD21 IS A MEMBER OF THE REGULATORS OF
CC       COMPLEMENT ACTIVATION (RCA) FAMILY.
CC   -!- SIMILARITY: CONTAINS 15 SUSHI (SCR) REPEATS.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD21 entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd21.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M26004; AAA35786.1; -.
DR   EMBL; M26016; AAB04638.1; -.
DR   EMBL; M24007; AAB04638.1; JOINED.
DR   EMBL; M24008; AAB04638.1; JOINED.
DR   EMBL; M24009; AAB04638.1; JOINED.
DR   EMBL; M24010; AAB04638.1; JOINED.
DR   EMBL; M24011; AAB04638.1; JOINED.
DR   EMBL; M26009; AAB04638.1; JOINED.
DR   EMBL; M26010; AAB04638.1; JOINED.
DR   EMBL; M26011; AAB04638.1; JOINED.
DR   EMBL; M26012; AAB04638.1; JOINED.
DR   EMBL; M26013; AAB04638.1; JOINED.
DR   EMBL; M26014; AAB04638.1; JOINED.
DR   EMBL; M26015; AAB04638.1; JOINED.
DR   PIR; A32036; A32036.
DR   PIR; A24319; A24319.
DR   PIR; B24319; B24319.
DR   PIR; C24319; C24319.
DR   PIR; D24319; D24319.
DR   PIR; E24319; E24319.
DR   PIR; F24319; F24319.
DR   HSSP; P10998; 1VVC.
DR   MIM; 120650; -.
DR   INTERPRO; IPR000436; -.
DR   PFAM; PF00084; sushi; 15.
KW   Complement pathway; Glycoprotein; Transmembrane; Repeat; Signal;
KW   Receptor; Sushi.
FT   SIGNAL        1     20
FT   CHAIN        21   1033       COMPLEMENT RECEPTOR TYPE 2.
FT   DOMAIN       21    971       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    972    999       POTENTIAL.
FT   DOMAIN     1000   1033       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       22    971       15 X SUSHI (SCR) REPEATS.
FT   REPEAT       22     83       SUSHI 1.
FT   REPEAT       90    147       SUSHI 2.
FT   REPEAT      153    211       SUSHI 3.
FT   REPEAT      214    272       SUSHI 4.
FT   REPEAT      275    343       SUSHI 5.
FT   REPEAT      350    407       SUSHI 6.
FT   REPEAT      409    467       SUSHI 7.
FT   REPEAT      470    523       SUSHI 8.
FT   REPEAT      526    594       SUSHI 9.
FT   REPEAT      601    658       SUSHI 10.
FT   REPEAT      661    715       SUSHI 11.
FT   REPEAT      718    780       SUSHI 12.
FT   REPEAT      787    844       SUSHI 13.
FT   REPEAT      850    908       SUSHI 14.
FT   REPEAT      911    969       SUSHI 15.
FT   DISULFID     23     65       BY SIMILARITY.
FT   DISULFID     51     82       BY SIMILARITY.
FT   DISULFID     91    132       BY SIMILARITY.
FT   DISULFID    118    146       BY SIMILARITY.
FT   DISULFID    154    197       BY SIMILARITY.
FT   DISULFID    183    210       BY SIMILARITY.
FT   DISULFID    215    256       BY SIMILARITY.
```

```
FT   DISULFID    242    271       BY SIMILARITY.
FT   DISULFID    276    325       BY SIMILARITY.
FT   DISULFID    305    342       BY SIMILARITY.
FT   DISULFID    351    393       BY SIMILARITY.
FT   DISULFID    379    406       BY SIMILARITY.
FT   DISULFID    410    453       BY SIMILARITY.
FT   DISULFID    439    466       BY SIMILARITY.
FT   DISULFID    471    509       BY SIMILARITY.
FT   DISULFID    495    522       BY SIMILARITY.
FT   DISULFID    527    576       BY SIMILARITY.
FT   DISULFID    556    593       BY SIMILARITY.
FT   DISULFID    602    644       BY SIMILARITY.
FT   DISULFID    630    657       BY SIMILARITY.
FT   DISULFID    662    699       BY SIMILARITY.
FT   DISULFID    685    714       BY SIMILARITY.
FT   DISULFID    719    762       BY SIMILARITY.
FT   DISULFID    748    779       BY SIMILARITY.
FT   DISULFID    788    830       BY SIMILARITY.
FT   DISULFID    816    843       BY SIMILARITY.
FT   DISULFID    851    894       BY SIMILARITY.
FT   DISULFID    880    907       BY SIMILARITY.
FT   DISULFID    912    955       BY SIMILARITY.
FT   DISULFID    941    968       BY SIMILARITY.
FT   CARBOHYD    121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    372    372       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    492    492       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    623    623       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    682    682       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    800    800       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    823    823       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    861    861       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    911    911       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    667    667       Q -> D (IN REF. 2).
FT   CONFLICT    902    902       Q -> G (IN REF. 2).
FT   CONFLICT    906    906       H -> L (IN REF. 2).
SQ   SEQUENCE   1033 AA;  112973 MW;  1749DB4A07847ADA CRC64;


  Query Match             8.5%;  Score 73.5;  DB 1;  Length 1033;
  Best Local Similarity   25.0%;  Pred. No. 8.8;
  Matches   32;  Conservative    8;  Mismatches   41;  Indels   47;  Gaps    5;

Qy      42 PPTNLRSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTG--LFGEEDVRFRSAPVYMPTV 99
           ||  |       : : :||    :|          |  |  | ||| ::  |  |: | |
Db     413 PPNILNGQKEDRHMVRFDPGTSIKY---------SCNPGYVLVGEESIQCTSEGVWTPPV 463

Qy     100 VLRRTPAC-AGGR---------------------------SVYTEAYVTIP--------V 123
           :  || | ||                             ||| |   |||
Db     464 PQCKVAACEATGRQLLTKPQHQFVRPDVNSSCGEGYKLSGSVYQECQGTIPWFMEIRLCK 523

Qy     124 GCTCVPEP 131
             || | |
Db     524 EITCPPPP 531


RESULT  12
PGCV_MOUSE
ID   PGCV_MOUSE     STANDARD;      PRT;  3358 AA.
AC   Q62059; Q62058;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   VERSICAN CORE PROTEIN PRECURSOR (LARGE FIBROBLAST PROTEOGLYCAN)
DE   (CHONDROITIN SULFATE PROTEOGLYCAN CORE PROTEIN 2) (PG-M).
GN   CSPG2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A. (VARIANTS V0; V1 AND V2).
RC   STRAIN=C57BL/6, AND SWISS WEBSTER; TISSUE=BRAIN;
RX   MEDLINE=95122551; PubMed=7822336;
RA   Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;
RT   "Multiple forms of mouse PG-M, a large chondroitin sulfate
RT   proteoglycan generated by alternative splicing.";
RL   J. Biol. Chem. 270:958-965(1995).
RN   [2]
RP   SEQUENCE OF 1-348 AND 3053-3358 FROM N.A. (VARIANT V3).
RC   STRAIN=C57BL/6;
RX   MEDLINE=95181355; PubMed=7876137;
RA   Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT   "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT   without a chondroitin sulfate attachment in region in mouse and human
RT   tissues.";
RL   J. Biol. Chem. 270:3914-3918(1995).
CC   -!- FUNCTION: MAY PLAY A ROLE IN INTERCELLULAR SIGNALING AND IN
CC       CONNECTING CELLS WITH THE EXTRACELLULAR MATRIX. MAY TAKE PART IN
CC       THE REGULATION OF CELL MOTILITY, GROWTH AND DIFFERENTIATION. BINDS
CC       HYALURONIC ACID.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; V0 (SHOWN HERE), V1,
CC       V2 AND V3; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: V2 IS FOUND ONLY IN BRAIN.
CC   -!- DEVELOPMENTAL STAGE: DISAPPEARS AFTER THE CARTILAGE DEVELOPMENT.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) REPEAT.
CC   -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D16263; BAA03796.1; -.
DR   EMBL; D28599; -; NOT_ANNOTATED_CDS.
DR   EMBL; D32040; BAA06802.1; -.
DR   HSSP; P00740; 1IXA.
DR   MGD; MGI:102889; CSPG2.
DR   INTERPRO; IPR000152; -.
DR   INTERPRO; IPR000436; -.
DR   INTERPRO; IPR000538; -.
DR   INTERPRO; IPR000561; -.
DR   INTERPRO; IPR001304; -.
DR   INTERPRO; IPR001438; -.
DR   INTERPRO; IPR001881; -.
DR   INTERPRO; IPR003006; -.
DR   PFAM; PF00008; EGF; 2.
DR   PFAM; PF00193; Xlink; 2.
DR   PFAM; PF00047; ig; 1.
DR   PFAM; PF00059; lectin_c; 1.
DR   PFAM; PF00084; sushi; 1.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS01241; LINK; 2.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   3358       VERSICAN CORE PROTEIN.
FT   DOMAIN       37    137       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      167    244       LINK 1.
FT   DOMAIN      265    334       LINK 2.
FT   DOMAIN      348   1308       GAG-ALPHA (GLUCOSAMINOGLYCAN ATTACHMENT
```

```
FT                                     DOMAIN).
FT   DOMAIN     1309   3052        GAG-BETA.
FT   DOMAIN     3052   3088        EGF-LIKE 1.
FT   DOMAIN     3090   3126        EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3129   3255        C-TYPE LECTIN.
FT   REPEAT     3258   3316        SUSHI.
FT   DISULFID     44    130        BY SIMILARITY.
FT   DISULFID    172    243        BY SIMILARITY.
FT   DISULFID    196    217        BY SIMILARITY.
FT   DISULFID    270    333        BY SIMILARITY.
FT   DISULFID    294    315        BY SIMILARITY.
FT   DISULFID   3056   3067        BY SIMILARITY.
FT   DISULFID   3061   3076        BY SIMILARITY.
FT   DISULFID   3078   3087        BY SIMILARITY.
FT   DISULFID   3094   3105        BY SIMILARITY.
FT   DISULFID   3099   3114        BY SIMILARITY.
FT   DISULFID   3116   3125        BY SIMILARITY.
FT   DISULFID   3132   3143        BY SIMILARITY.
FT   DISULFID   3160   3252        BY SIMILARITY.
FT   DISULFID   3228   3244        BY SIMILARITY.
FT   DISULFID   3259   3302        BY SIMILARITY.
FT   DISULFID   3288   3315        BY SIMILARITY.
FT   CARBOHYD     57     57        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    351    351        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    441    441        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    807    807        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    914    914        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    951    951        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1305   1305        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1372   1372        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1679   1679        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2054   2054        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2244   2244        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2362   2362        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2627   2627        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3030   3030        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3332   3332        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3342   3342        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    349   1308        MISSING (IN ISOFORM V1).
FT   VARSPLIC   1309   3052        MISSING (IN ISOFORM V2).
FT   VARSPLIC    349   3052        MISSING (IN ISOFORM V3).
FT   CONFLICT    348    348        P -> R (IN REF. 2).
SQ   SEQUENCE   3358 AA;  366938 MW;  071B80026BC0762D CRC64;


  Query Match             8.4%;  Score 73;  DB 1;  Length 3358;
  Best Local Similarity   32.1%;  Pred. No. 38;
  Matches   26;  Conservative   10;  Mismatches   33;  Indels   12;  Gaps    4;

Qy      81 LFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVY-TE-AYVTIPVGCTCVP-----EPEK 133
           | |  :       ||:|   | :|  |  | : | || :||     ||| |     : |
Db    3035 LIGISEESVEGTAVYLPGPDLCKTNPCLNGGTCYPTETSYV-----CTCAPGYSGDQCEL 3089

Qy     134 DADSINSSIDKQGAKLLLGPN 154
           | |  :|:  : ||  : | |
Db    3090 DFDECHSNPCRNGATCVDGFN 3110


RESULT  13
CDN1_HUMAN
ID   CDN1_HUMAN     STANDARD;      PRT;   164 AA.
AC   P38936;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CYCLIN-DEPENDENT KINASE INHIBITOR 1 (MELANOMA DIFFERENTIATION
DE   ASSOCIATED PROTEIN 6) (MDA-6) (P21) (CDK-INTERACTING PROTEIN 1).
GN   CDKN1A OR CDKN1 OR CIP1 OR WAF1 OR MDA6 OR SDI1 OR PIC1 OR CAP20.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94061996; PubMed=8242751;
RA   Harper J.W., Adami G.R., Wei N., Keyomarsi K., Elledge S.J.;
RT   "The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1
RT   cyclin-dependent kinases.";
RL   Cell 75:805-816(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94061997; PubMed=8242752;
RA   El-Deiry W.S., Tokino T., Velculescu V.E., Levy D.B., Parsons R.,
RA   Trent J.M., Lin D., Mercer W.E., Kinzler K.W., Vogelstein B.;
RT   "WAF1, a potential mediator of p53 tumor suppression.";
RL   Cell 75:817-825(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94081955; PubMed=8259214;
RA   Xiong Y., Hannon G.J., Zhang H., Casso D., Kobayashi R., Beach D.;
RT   "p21 is a universal inhibitor of cyclin kinases.";
RL   Nature 366:701-704(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Jiang H., Fisher P.B.;
RT   "Use of a sensitive and efficient subtraction hybridization protocol
RT   for the identification of genes differentially regulated during the
RT   induction of differentiation in human melanoma cells.";
RL   Mol. Cell. Differ. 1:285-299(1993).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Jiang H., Lin J., Herlyn M., Kerbel R.S., Weissman B.E.,
RA   Welch D.R., Fisher P.B.;
RL   Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94170884; PubMed=8125163;
RA   Noda A., Ning Y., Venable S.F., Pereira-Smith O.M., Smith J.R.;
RT   "Cloning of senescent cell-derived inhibitors of DNA synthesis using
RT   an expression screen.";
RL   Exp. Cell Res. 211:90-98(1994).
RN   [7]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95384154; PubMed=7655464;
RA   Mousses S., Oezcelik H., Lee P.D., Malkin D., Bull S.B.,
RA   Andrulis I.L.;
RT   "Two variants of the CIP1/WAF1 gene occur together and are associated
RT   with human cancer.";
RL   Hum. Mol. Genet. 4:1089-1092(1995).
RN   [8]
RP   SEQUENCE FROM N.A.
RA   Palmer S.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 139-160.
RX   MEDLINE=97015085; PubMed=8861913;
RA   Gulbis J.M., Kelman Z., Hurwitz J., O'Donnell M., Kuriyan J.;
RT   "Structure of the C-terminal region of p21(WAF1/CIP1) complexed with
RT   human PCNA.";
RL   Cell 87:297-306(1996).
CC   -!- FUNCTION: MAY BE THE IMPORTANT INTERMEDIATE BY WHICH P53 MEDIATES
CC       ITS ROLE AS AN INHIBITOR OF CELLULAR PROLIFERATION IN RESPONSE TO
CC       DNA DAMAGE. MAY BIND TO AND INHIBIT CYCLIN-DEPENDENT KINASE
CC       ACTIVITY, PREVENTING PHOSPHORYLATION OF CRITICAL CYCLIN-DEPENDENT
CC       KINASE SUBSTRATES AND BLOCKING CELL CYCLE PROGRESSION.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: IS EXPRESSED IN ALL ADULT HUMAN TISSUES,
CC       WITH 5-FOLD LOWER LEVELS OBSERVED IN THE BRAIN.
CC   -!- INDUCTION: BY THE P53 TUMOR SUPPRESSOR/ONCOGENE. ALSO BY MEZEREIN
CC       (ANTILEUKEMIC COMPOUND) AND BY IFN-BETA.
CC   -!- SIMILARITY: THE N-TERMINAL OF CIP1 AND KIP ARE SIMILAR.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; L25610; AAA16109.1; -.
DR   EMBL; S67388; AAB29246.1; -.
DR   EMBL; U09579; AAA85641.1; -.
DR   EMBL; U03106; AAC04313.1; -.
DR   EMBL; L26165; AAA19811.1; -.
DR   EMBL; L47233; AAB59560.1; ALT_INIT.
DR   EMBL; Z85996; CAB06656.1; -.
DR   PIR; S39357; S39357.
DR   SWISS-2DPAGE; P38936; HUMAN.
DR   MIM; 116899; -.
KW   Cell cycle; Nuclear protein; Zinc-finger.
FT   ZN_FING      13     41       C4-TYPE (POTENTIAL).
FT   DOMAIN      141    156       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   164 AA;  18119 MW;  98D1E7C519ADFCA9 CRC64;


  Query Match             8.3%;  Score 71.5;  DB 1;  Length 164;
  Best Local Similarity   25.2%;  Pred. No. 1.7;
  Matches   39;  Conservative   10;  Mismatches   53;  Indels   53;  Gaps    6;

Qy      30 PAG-------GRPADRRFRPPTNLRSVSPWAYRISYDPARYPRYLPEAYCLCRGC----- 77
           |||       |  | ||   | :   :|                :   |  ||
Db       4 PAGDVRQNPCGSKACRRLFGPVDSEQLS----------------RDCDALMAGCIQEAR 46

Qy      78 ------------LTGLFGEEDVRFRSAP-VYMPTVVLRRTPACAGGRSVYT--------- 115
                       | | |  | ||    | :|:||   |      |||   |
Db      47 ERWNFDFVTETPLEGDFAWERVRGLGLPKLYLPTGPRRGRDELGGGRRPGTSPALLQGTA 106

Qy     116 -EAYVTIPVGCTCVPEPEKDAD-SINSSIDKQGAK 148
            | :| : : || ||   : |: |     | || |
Db     107 EEDHVDLSLSCTLVPRSGEQAEGSPGGPGDSQGRK 141


RESULT  14
YM40_MYCTU
ID   YM40_MYCTU     STANDARD;      PRT;   265 AA.
AC   Q10522;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL 27.5 KDA PROTEIN RV2240C.
GN   RV2240C OR MTCY427.21C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA   Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; Z70692; CAA94661.1; -.
DR   TUBERCULIST; Rv2240c; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM     89    109       POTENTIAL.
SQ   SEQUENCE   265 AA;  27529 MW;  F9E7ACF3A736B90A CRC64;


  Query Match             8.2%;  Score 70.5;  DB 1;  Length 265;
  Best Local Similarity   25.4%;  Pred. No. 3.8;
  Matches   44;  Conservative   15;  Mismatches   69;  Indels   45;  Gaps    8;

Qy      28 SCPAGGRP---ADRRFRPPTNLRSV-----SPWAYRISYD--------------PARYPR 65
           :|   |||     || |    ||||     :|   |   |              | |:
Db      24 ACCLDGRPPIVPHRRRRRIAALRSVLRMRDTPRPARSRCDQVTSHAVLIGWRAVPRRHGG 83

Qy      66 YLPEAYCLCRGC----LTGLFGEEDVR-------FRSAPVYMPTVVLRRTPACAGG---- 110
             ||    |  ||    | |: |   |           || |  :|    : : |
Db      84 ELPRRGALALGCIALLLMGIVGCTTVTDGTAMPDTNVAPAYRSSVSASVSASAATSSIRE 143

Qy     111 ----RSVYTEAYVTIPVGCTCVPEPEKDA-DSINSSIDKQGAKLLLGPNDAPA 158
               :|: |:|   |   |  :    ||| | :|: :         || : ||
Db     144 SQRQQSLTTKA---IRTSCDALAATSKDAIDKVNAYVAAFNQGRNTGPTEGPA 193


RESULT  15
GLYC_NEUCR
ID   GLYC_NEUCR     STANDARD;      PRT;   479 AA.
AC   P34898;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SERINE HYDROXYMETHYLTRANSFERASE, CYTOSOLIC (EC 2.1.2.1) (SERINE
DE   METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
GN   FOR.
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92195285; PubMed=1532227;
RA   McClung C.R., Davis C.R., Page K.M., Denome S.A.;
RT   "Characterization of the formate (for) locus, which encodes the
RT   cytosolic serine hydroxymethyltransferase of Neurospora crassa.";
RL   Mol. Cell. Biol. 12:1412-1421(1992).
CC   -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC   -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC       H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC       HORMONES AND OTHER COMPONENTS.
CC   -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYMES: A
CC       CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.
CC   -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; M81918; CAB26403.1; -.
DR   PIR; A42241; A42241.
DR   INTERPRO; IPR001085; -.
DR   PFAM; PF00464; SHMT; 1.
DR   PROSITE; PS00096; SHMT; 1.
KW   Transferase; Pyridoxal phosphate; One-carbon metabolism.
FT   BINDING     249    249       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   479 AA;  52821 MW;  06E196B0395DE1E4 CRC64;
```

```
  Query Match             8.2%;  Score 70.5;  DB 1;  Length 479;
  Best Local Similarity   19.7%;  Pred. No. 7.5;
  Matches   38;  Conservative   23;  Mismatches   59;  Indels   73;  Gaps    7;

Qy       9 LSAFHHTLQLGPREQARNASC-------------------------PAGGRPADRRFRP 42
           | |||    | |::   |  |                         | ||  :     |
Db      93 LEAFH----LDPKQWGVNVQCLSGSPANLQVYQAIMPVHGRLMGLDLPHGGHLSHGYQTP 148

Qy      43 PTNLRSVSPW----AYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPT 98
             : :|| :     ||:: |                  |||  : |   ::| :: |
Db     149 QRKISAVSTYFETMPYRVNID-----------------TGLI-DYDTLEKNAQLFRPK 188

Qy      99 VVLRRTPA--------------CAGGRSVYTEAYVTIPVGCTCVPEPEKDADSINSSID 143
           |::  | |                |   |   |:::  :    :| |   || : ::
Db     189 VLVAGTSAYCRLIDYERMRKIADSVGAYLVVDMAHISGLIASEVIPSPFLYADVVTTTH 248

Qy     144 KQGAKLLLGPNDA 156
           |     | ||  |
Db     249 KS----LRGPRGA 257


Search completed: February 16, 2001, 12:53:01
Job time: 191 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:52:39 ; Search time 38.3 Seconds
(without alignments)
489.641 Million cell updates/sec

Title: US-09-320-713-4
Perfect score: 864
Sequence: 1 NSARARAVLSAFHHTLQLGP.....SIDKQGAKLLLGPNDAPAGP 160

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	192	22.2	197	4	Q9P0M4	Q9p0m4 homo sapien
2	177	20.5	180	11	Q9QXT6	Q9qxt6 mus musculu
3	174	20.1	180	4	Q9UHF5	Q9uhf5 homo sapien
4	119.5	13.8	78	4	Q9NUE6	Q9nue6 homo sapien
5	107.5	12.4	151	12	O40633	O40633 saimiriine
6	99	11.5	148	5	Q19778	Q19778 caenorhabdi
7	84.5	9.8	500	4	Q16753	Q16753 homo sapien
8	84.5	9.8	552	4	O95723	O95723 homo sapien
9	83.5	9.7	226	5	Q9XUH9	Q9xuh9 caenorhabdi
10	80.5	9.3	805	5	Q9VFD4	Q9vfd4 drosophila
11	79.5	9.2	520	4	Q9UBG7	Q9ubg7 homo sapien
12	77.5	9.0	1045	6	O46545	O46545 ovis aries
13	77	8.9	354	5	Q9XWJ4	Q9xwj4 caenorhabdi
14	76.5	8.9	733	5	Q9VTX9	Q9vtx9 drosophila
15	76	8.8	242	2	Q9PL24	Q9pl24 chlamydia m
16	76	8.8	1403	12	Q9QP21	Q9qp21 gallid herp
17	75.5	8.7	425	4	O75038	O75038 homo sapien
18	75.5	8.7	509	10	O48780	O48780 arabidopsis
19	75	8.7	178	2	Q9X0U1	Q9x0u1 thermotoga
20	75	8.7	242	2	O84019	O84019 chlamydia t
21	74.5	8.6	504	2	P72236	P72236 pseudomonas
22	74.5	8.6	1887	5	Q9VYX6	Q9vyx6 drosophila
23	74	8.6	370	11	O08771	O08771 rattus norv
24	74	8.6	495	2	Q50140	Q50140 mycobacteri
25	74	8.6	1006	5	Q9VBP9	Q9vbp9 drosophila
26	74	8.6	2390	11	O08592	O08592 rattus norv
27	73.5	8.5	195	12	Q87073	Q87073 pseudorabie
28	73.5	8.5	577	12	Q9Q3F7	Q9q3f7 pseudorabie
29	73.5	8.5	1087	4	Q14212	Q14212 homo sapien
30	73	8.4	353	10	Q9LHG4	Q9lhg4 arabidopsis
31	73	8.4	903	4	Q9UPX1	Q9upx1 homo sapien
32	72.5	8.4	578	12	Q9PYC0	Q9pyc0 pseudorabie
33	72.5	8.4	2087	11	Q9WUE8	Q9wue8 rattus norv
34	72.5	8.4	2158	11	Q9WU13	Q9wu13 rattus norv
35	72.5	8.4	2167	11	Q9WV48	Q9wv48 rattus norv
36	72	8.3	685	4	O15271	O15271 homo sapien
37	72	8.3	1223	10	Q9SNI5	Q9sni5 oryza sativ
38	71.5	8.3	516	4	Q9ULV9	Q9ulv9 homo sapien
39	71.5	8.3	557	2	Q9ZFY6	Q9zfy6 thermus aqu
40	71.5	8.3	762	4	O95914	O95914 homo sapien
41	71.5	8.3	789	4	O75626	O75626 homo sapien
42	71.5	8.3	2129	12	Q9J6K9	Q9j6k9 rubella vir
43	71	8.2	464	1	Q9YEB8	Q9yeb8 aeropyrum p
44	71	8.2	1277	12	Q08547	Q08547 cercopithec
45	71	8.2	1279	12	Q66031	Q66031 cercopithec

ALIGNMENTS

RESULT 1
Q9P0M4
ID Q9P0M4 PRELIMINARY; PRT; 197 AA.
AC Q9P0M4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE INTERLEUKIN 17C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li H., Chen J., Huang A., Stinson J., Heldens S., Foster J., Dowd P.,
RA Gurney A.L., Wood W.I.;
RT "Cloning and Characterization of IL-17B and IL17C, Two New Members of
RT the IL-17 Cytokine Family.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:773-778(2000).
DR EMBL; AF152099; AAF28105.1; -.
SQ SEQUENCE 197 AA; 21765 MW; BAE0152E18DE7D08 CRC64;

Query Match 22.2%; Score 192; DB 4; Length 197;
Best Local Similarity 34.6%; Pred. No. 1.9e-12;
Matches 45; Conservative 17; Mismatches 58; Indels 10; Gaps 4;

```
Qy       7 AVLSAFHHTLQLGPREQ-ARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPR 65
           |::|:        |  |: :    ||    || :      |: ||:||| ||:  |  |||:
Db      65 ALVSSLEAASHRGRHERPSATTQCPV-LRP-EEVLEADTHQRSISPWRYRVDTDEDRYPQ 122

Qy      66 YLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEA-----Y 118
             |  | ||||||:    | |    |  :    :|||| |    | : |        :
Db     123 KLAFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRPCSRDGSGLPTPGAFAFHTEF 182

Qy     119 VTIPVGCTCV 128
           : :|||||||
Db     183 IHVPVGCTCV 192
```

RESULT 2
Q9QXT6

```
ID   Q9QXT6      PRELIMINARY;      PRT;   180 AA.
AC   Q9QXT6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CYTOKINE-LIKE PROTEIN ZCYTO7.
GN   ZCYTO7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Presnell S., Gilbert T., Whitmore T., Foster D., Hart C., Lehner J.,
RA   Martinez T., Hoffman R., O'Hara P.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF184970; AAF01319.1; -.
SQ   SEQUENCE   180 AA;  20309 MW;  E26F4C72001997C5 CRC64;


  Query Match             20.5%;  Score 177;  DB 11;  Length 180;
  Best Local Similarity   40.2%;  Pred. No. 6.2e-11;
  Matches   45;  Conservative   11;  Mismatches   46;  Indels   10;  Gaps    2;

Qy      23 QARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLF 82
           | ||:| ||  :          :| ||:||| | |::||:| |  |||| ||| ||:
Db      72 QLRNSSEPAKKKCEVNLQLWLSNKRSLSPWGYSINHDPSRIPADLPEARCLCLGCVNPFT 131

Qy      83 GEEDVRFRSAPVYMPTVVLRR------TPACAGGRSVYTEAYVTIPVGCTCV 128
             :||    | ||:    | ||       |     | |      || |||||:
Db     132 MQEDRSMVSVPVFSQVPVRRRLCPQPPRPGPCRQRVVME----TIAVGCTCI 179


RESULT   3
Q9UHF5
ID   Q9UHF5      PRELIMINARY;      PRT;   180 AA.
AC   Q9UHF5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CYTOKINE-LIKE PROTEIN ZCYTO7 (INTERLEUKIN 20) (INTERLEUKIN 17B).
GN   ZCYTO7 OR IL20.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Presnell S., Gilbert T., Whitmore T., Foster D., Hart C., Lehner J.,
RA   Martinez T., Hoffman R., O'Hara P.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Shi Y., Ullrich S.J., Zhang J., Connolly K., Grzegorzewski K.J.,
RA   Barber M.C., Wang W., Wathen K., Hodge V., Fisher C.L., Olsen H.,
RA   Ruben S.M., Knyazev I., Cho Y.H., Kao V., Wilkinson K.A.,
RA   Carrell J.A., Ebner R.;
RT   "A novel cytokine receptor-ligand pair. Identification, molecular
RT   characterization, and in vivo immunomodulatory activity.";
RL   J. Biol. Chem. 275:19167-19176(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Li H., Chen J., Huang A., Stinson J., Heldens S., Foster J., Dowd P.,
RA   Gurney A.L., Wood W.I.;
RT   "Cloning and Characterization of IL-17B and IL17C, Two New Members of
RT   the IL-17 Cytokine Family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:773-778(2000).
DR   EMBL; AF184969; AAF01318.1; -.
DR   EMBL; AF212311; AAF78775.1; -.
DR   EMBL; AF152098; AAF28104.1; -.
SQ   SEQUENCE   180 AA;  20437 MW;  F1B0BC1446D0B14A CRC64;


  Query Match             20.1%;  Score 174;  DB 4;  Length 180;
  Best Local Similarity   43.7%;  Pred. No. 1.3e-10;
  Matches   38;  Conservative    9;  Mismatches   38;  Indels    2;  Gaps    1;

Qy      44 TNLRSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRR 103
           :| ||:||| | |::||:| |  |||| ||| ||:     :||    | ||:    | ||
Db      93 SNKRSLSPWGYSINHDPSRIPVDLPEARCLCLGCVNPFTMQEDRSMVSVPVFSQVPVRRR 152

Qy     104 --TPACAGGRSVYTEAYVTIPVGCTCV 128
             |    |         || |||||:
Db     153 LCPPPPRTGPCRQRAVMETIAVGCTCI 179


RESULT   4
Q9NUE6
ID   Q9NUE6      PRELIMINARY;      PRT;    78 AA.
AC   Q9NUE6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   DJ108C2.3 (PUTATIVE NOVEL PROTEIN SIMILAR TO IL17 (INTERLEUKIN 17
DE   (CYTOTOXIC T-LYMPHOCYTE-ASSOCIATED SERINE ESTERASE 8)) (CYTOTOXIC T
DE   LYMPHOCYTE-ASSOCIATED ANTIGEN 8, CTLA8)) (FRAGMENT).
GN   DJ108C2.3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Almeida J.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL034343; CAB75300.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   78 AA;  8704 MW;  66C3C34E7ACC2790 CRC64;


  Query Match             13.8%;  Score 119.5;  DB 4;  Length 78;
  Best Local Similarity   37.8%;  Pred. No. 2.5e-05;
  Matches   28;  Conservative   13;  Mismatches   28;  Indels    5;  Gaps    3;

Qy      56 ISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSV-Y 114
           :::|| |||  : :| |   ||:    |:||:   | |:   |:|:||      | || :
Db       1 VTWDPNRYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQQETLVVRRKHQ---GCSVSF 56

Qy     115 TEAYVTIPVGCTCV 128
             | : ||||||
Db      57 QLEKVLVTVGCTCV 70


RESULT   5
O40633
ID   O40633      PRELIMINARY;      PRT;   151 AA.
AC   O40633;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   INTERLEUKIN 17.
OS   Saimiriine herpesvirus 2.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=10381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-488;
RX   MEDLINE=98037620; PubMed=9371569;
RA   Knappe A., Hiller C., Thurau M., Wittmann S., Hofmann H.,
RA   Fleckenstein B., Fickenscher H.;
RT   "The superantigen-homologous viral immediate-early gene ie14/vsag in
RT   herpesvirus saimiri-transformed human T cells.";
RL   J. Virol. 71:9124-9133(1997).
DR   EMBL; Y13183; CAA73627.1; -.
SQ   SEQUENCE   151 AA;  17189 MW;  49B2C9430C46BE32 CRC64;
```

```
  Query Match             12.4%;  Score 107.5;  DB 12;  Length 151;
  Best Local Similarity   31.4%;  Pred. No. 0.00093;
  Matches   27;  Conservative   12;  Mismatches   36;  Indels   11;  Gaps    3;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPA 106
           || |||    : |  |||  : || |   ||:    |  |    | |:    :|:|:
Db      65 RSTSPWTLYRNEDQDRYPSVIWEAKCRYLGCVNA-DGNVDYHMNSVPIQQEILVVRK--- 120

Qy     107 CAGGRSVYTEAY----VTIPVGCTCV 128
             | :    ::     : : ||||||
Db     121 ---GHNPCPNSFRLEKMLVTVGCTCV 143


RESULT   6
Q19778
ID   Q19778      PRELIMINARY;      PRT;   148 AA.
AC   Q19778;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   F25D1.3 PROTEIN.
GN   F25D1.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kelly P.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z73973; CAA98268.1; -.
SQ   SEQUENCE   148 AA;  16738 MW;  ADB51F438DFB3940 CRC64;


  Query Match             11.5%;  Score 99;  DB 5;  Length 148;
  Best Local Similarity   29.8%;  Pred. No. 0.007;
  Matches   28;  Conservative   17;  Mismatches   35;  Indels   14;  Gaps    5;

Qy      47 RSVSPWAYRISYDPARYPRYLPEAYCLCRGC--LTGLFGEEDVRFRSAPVYMPTVVLRRT 104
           |:: ||  |::|  :| |: : |: |||| || |          |:     :|||
Db      52 RALCPWDSRVNYQESREPKLIAESVCLCRKSRGSTGAF--------CMPIVRKVPILRRV 103

Qy     105 PAC--AGGRSVYTEAYVTIPVGC-TCVPEPEKDA 135
           :|  : |   |  :   | ||| : :|  :: |
Db     104 -SCDRSTGLWNYVRSTELITVGCHSVLPRTQRAA 136


RESULT   7
Q16753
ID   Q16753      PRELIMINARY;      PRT;   500 AA.
AC   Q16753;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DEBRISOQUINE 4-HYDROXYLASE MUTANT ALLELE (EC 1.14.14.1).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gonzalez F.J.;
RL   Submitted (JUL-1990) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   EMBL; M33189; AAA35737.1; -.
DR   INTERPRO; IPR001128; -.
DR   INTERPRO; IPR002401; -.
DR   PFAM; PF00067; p450; 2.
DR   PRINTS; PR00385; P450.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.
FT   BINDING     446    446       HEME (BY SIMILARITY).
SQ   SEQUENCE   500 AA;  55889 MW;  D5293E9BF74692C8 CRC64;


  Query Match              9.8%;  Score 84.5;  DB 4;  Length 500;
  Best Local Similarity   30.8%;  Pred. No. 0.87;
  Matches   41;  Conservative   12;  Mismatches   53;  Indels   27;  Gaps    6;

Qy       3 ARARAVLSAFHHTLQLGPREQARNASCPAGGR--------PADRRFRP-------PTNLR 47
             ||| |:|    |  |||| ::   |    |||         |  | |||        :|:
Db     132 ARAEALLRL--HLAQLGPGQEVAGAVGDRGGRLPLCRLRQPLRRPFRPNGLLDKAVSNVI 189

Qy      48 SVSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPAC 107
           :      |  ||  |: | |  |          ||  ||    |      :|  ||   ||
Db     190 ASLTCGRRFEYDDPRFLRLLDLAQ-------EGL-KEESGFLREVLNAVP--VLLHIPAL 239

Qy     108 AGGRSVYTEAYVT 120
           ||    : :|::|
Db     240 AGKVLRFQKAFLT 252


RESULT   8
O95723
ID   O95723      PRELIMINARY;      PRT;   552 AA.
AC   O95723;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   DJ453C12.2 (SIMILAR TO TRANSCRIPTION FACTOR RPB-L).
GN   DJ453C12.2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bridgeman A.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL021578; CAA16521.1; -.
SQ   SEQUENCE   552 AA;  60422 MW;  5E2EC4B12237AF43 CRC64;


  Query Match              9.8%;  Score 84.5;  DB 4;  Length 552;
  Best Local Similarity   28.1%;  Pred. No. 0.97;
  Matches   39;  Conservative   11;  Mismatches   44;  Indels   45;  Gaps    8;

Qy      18 LGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPRYL----PEAYCL 73
           ||| || :       |  |      |     |||          | || |    |:
Db     436 LGPGEQGK-------GVHASSESPQP------SPWC-----STPRSPRSLVCVVPDVAAF 477

Qy      74 CRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGG-------RSVYTEAYVTIPVGCT 126
           |           | |:  ||: :|  ::|     | |              ||  |   | |
Db     478 C----------SDWRWLRAPITIPMSLVR-----ADGLFYPSAFSFTYTPEYSVRP-GHP 521

Qy     127 CVPEPEKDADSINSSIDKQ 145
            ||||  |||::  || ::
```

```
Db    522 GVPEPATDADALLESIHQE 540


RESULT   9
Q9XUH9
ID   Q9XUH9      PRELIMINARY;      PRT;   226 AA.
AC   Q9XUH9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ZK39.8 PROTEIN.
GN   ZK39.8.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Kershaw J.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z82093; CAB05021.1; -.
DR   INTERPRO; IPR001304; -.
DR   PFAM; PF00059; lectin_c; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   226 AA;  24150 MW;  DB4C40BFFF904200 CRC64;


  Query Match             9.7%;  Score 83.5;  DB 5;  Length 226;
  Best Local Similarity   25.6%;  Pred. No. 0.47;
  Matches   32;  Conservative   18;  Mismatches   34;  Indels   41;  Gaps    8;

Qy      26 NASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARYPRYLPEAYCLCRGC-LTGLFGE 84
           |  | ||      ||  |:       |  |:   |  | : | |: |  :|  |||:  :
Db      52 NRGCDAGW----TRFNRPS-----GGWCVRVF--PGTYHQPLAESRCQSQGAVLTGVQNQ 100

Qy      85 EDVRFRSAPVYMPTVV---------LRRTPACAGGRSVYTEAYVTIPVGCTCVPEPEKDA 135
           |: : : | : :| :          | |||||:             |:  :|
Db     101 EEAK-KIASLLLPQISQQSGSIYIGLHRTPACSKS----------PISSSC-------- 140

Qy     136 DSINS 140
           :|:||
Db     141 NSMNS 145


RESULT  10
Q9VFD4
ID   Q9VFD4      PRELIMINARY;      PRT;   805 AA.
AC   Q9VFD4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG18442 PROTEIN.
GN   CG18442.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003707; AAF55126.1; -.
DR   FLYBASE; FBgn0038287; CG18442.
SQ   SEQUENCE   805 AA;  86321 MW;  FDF9EFEF14E69957 CRC64;


  Query Match             9.3%;  Score 80.5;  DB 5;  Length 805;
  Best Local Similarity   21.5%;  Pred. No. 3.8;
  Matches   42;  Conservative   24;  Mismatches   60;  Indels   69;  Gaps   10;

Qy       1 NSARAR-----AVLSAFHHTLQLGP--------REQARNASCPAGGR---PADRRFRPPT 44
           || :|:      :| :  : |:  |           ::: : ||   :   |     :||
Db     434 NSVQAKREFFQSVANGNNGTVSHSPHIIEYATVERKSKFSCCPTSKKEEQPGQEELQPPP 493

Qy      45 NLRSVSPWAYRISYDPARYP-------RYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYM- 96
             :|   ::|  ||| |          || |               :: ||  ::||
Db     494 R----TPTTEQLSPPPARPPKSAELLQRYSP--------------KKQVRIAASPVMQP 534

Qy      97 --------PTVVLRRTPACAGGRSVYTEAY------VTIPVGCTCVPEPEKDADSINSSI 142
                 | :  | :|   | :|  | |          |: |   | |    :| |||
Db     535 QERRELCPPQLPPRGSPTLDGSQSSPTNAVSGPKKPPLPPIACR--PRPSNGVNSPNSS- 591

Qy     143 DKQGAKLLLGPNDAP 157
                   |  ||
Db     592 ---------SPGSAP 597


RESULT  11
```

```
Q9UBG7
ID   Q9UBG7      PRELIMINARY;      PRT;   520 AA.
AC   Q9UBG7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   TRANSCRIPTION FACTOR RBP-L.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PANCREAS;
RA   Koyama K., Isaka S., Okamura S.;
RT   "112-pan.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Isaka S., Koyama K., Nakamura Y., Okamura S., Azuma C., Kimura T.;
RT   "Human RBP-L.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB024964; BAA88232.1; -.
DR   EMBL; AB026048; BAA86121.1; -.
SQ   SEQUENCE   520 AA;  57038 MW;  BACF92459A77ACFB CRC64;

  Query Match             9.2%;  Score 79.5;  DB 4;  Length 520;
  Best Local Similarity   28.3%;  Pred. No. 3;
  Matches   28;  Conservative   12;  Mismatches   36;  Indels   23;  Gaps    4;

Qy      54 YRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDVRFRSAPVYMPTVVLRRTPACAGG--- 110
           ||   :||     :|:    |           | |:  ||: :|  ::|     | |
Db     426 YRYGVEPAVLVCVVPDVAAFC----------SDWRWLRAPITIPMSLVR-----ADGLFY 470

Qy     111 ----RSVYTEAYVTIPVGCTCVPEPEKDADSINSSIDKQ 145
                 ||  |   | |   ||||  |||::  || ::
Db     471 PSAFSFTYTPEYSVRP-GHPGVPEPATDADALLESIHQE 508


RESULT  12
O46545
ID   O46545      PRELIMINARY;      PRT;  1045 AA.
AC   O46545;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   COMPLEMENT RECEPTOR TYPE 2.
GN   CR2.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE ALPINE;
RA   Hein W.R., Dudler L., Marston W., Landsverk T., Young A., Avila D.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF038131; AAB92375.1; -.
DR   HSSP; P10998; 1VVD.
DR   INTERPRO; IPR000436; -.
DR   PFAM; PF00084; sushi; 14.
SQ   SEQUENCE   1045 AA;  115561 MW;  FF58E1A2892C0D59 CRC64;

  Query Match             9.0%;  Score 77.5;  DB 6;  Length 1045;
  Best Local Similarity   25.9%;  Pred. No. 10;
  Matches   29;  Conservative   13;  Mismatches   43;  Indels   27;  Gaps    5;

Qy      42 PPTNLRSVSPWAYRISYDPARYPRYLPEAYCLCRGCLTG--LFGEEDVRFRSAPVYMPTV 99
           ||  |       :|: :||    ||          |  |  | ||| :|     |::||
Db     423 PPKILNGQKEDRHRVRFDPGTSIRY---------SCDPGYVLVGEESIRCTPDGVWIPTA 473

Qy     100 VLRRTPACAGGRSVYTEAYVTIPVGCTCVPEPEKD--ADSINSSIDKQGAKL 149
           : :   |              |||     :|:       :||| | :| :|
Db     474 PICKAAECE------------PVGKQVFKKPKNQFIRPDVNSSCD-EGYRL 511


RESULT  13
Q9XWJ4
ID   Q9XWJ4      PRELIMINARY;      PRT;   354 AA.
AC   Q9XWJ4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Y51H1A.7 PROTEIN.
GN   Y51H1A.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Smye R.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; AL032644; CAA21670.1; -.
DR   INTERPRO; IPR002965; -.
DR   PRINTS; PR01217; PRICHEXTENSN.
SQ   SEQUENCE   354 AA;  40127 MW;  6F114E6450DAD63E CRC64;

  Query Match             8.9%;  Score 77;  DB 5;  Length 354;
  Best Local Similarity   25.0%;  Pred. No. 3.6;
  Matches   36;  Conservative   20;  Mismatches   56;  Indels   32;  Gaps    8;

Qy      20 PREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPARY-PRYLPEAYCLCRGCL 78
           |  : | :: ||   ||  ::||    ||   :  ||      | ||  |:|| | :
Db     141 PERRPRRSAAPAS--PAKPQYRPQPPARSHEEYTQRIMNKVKEYMPRRRPDAYDLQQ--- 195

Qy      79 TGLFGEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYTEAYVTIPVGCTCVPEPEKDADSI 138
                |::  | :    .:    | | |       :  |     |   |  | |::
Db     196 ----GQQQSRLQ----MLREQTLLRAP-------LEEEK----PPVATMAPVPDE----- 231

Qy     139 NSSI--DKQGAKLLLGPNDAPAGP 160
           : :|   : | |  :|| :|   |
Db     232 HKAIPWTRPGRKRAMGPGEAATAP 255


RESULT  14
Q9VTX9
ID   Q9VTX9      PRELIMINARY;      PRT;   733 AA.
AC   Q9VTX9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CG10663 PROTEIN.
```

```
GN   CG10663.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003541; AAF49916.1; -.
DR   HSSP; P00750; 1RTF.
DR   FLYBASE; FBgn0036287; CG10663.
DR   INTERPRO; IPR000884; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001314; -.
DR   INTERPRO; IPR002465; -.
DR   INTERPRO; IPR002965; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR01217; PRICHEXTENSN.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
DR   PROSITE; PS00340; RECEPTOR_CYTOKINES_2; UNKNOWN_1.
SQ   SEQUENCE   733 AA;  84355 MW;  8E47F6A30E8CB61B CRC64;
```

```
  Query Match             8.9%;  Score 76.5;  DB 5;  Length 733;
  Best Local Similarity   26.1%;  Pred. No. 9.1;
  Matches   36;  Conservative   15;  Mismatches   42;  Indels   45;  Gaps    8;

Qy      26 NASCPAGGRP-ADRR---FRPPTNLRSVSPWAYRISYD---PARYPRYLPEAYCLCRGCL 78
           ||:  |   |  | |   ::|| |  |: |  :   |    || :|   |  |
Db     134 NAAAAASSGPTVDYRYTEYKPPPNYASLYPSTHPPIYTPTYPAAHPPTQPPTY------- 186

Qy      79 TGLFGEEDVRFRSAPVYMPT----------VVLRRTPACAGGRSVYTEAYVTIPVGCTCV 128
                     | : ||           |:|  |    |: :::|  |   :
Db     187 -----------PPPTHPPTPPPTPPTNPDEVVRHLPKYP-GKDIFSEEDVEDEL----- 228

Qy     129 PEPEKDADSINSSIDKQG 146
           ||| | | |:|  |:||
Db     229 -EPEPDEDLISS--DEQG 243
```

```
RESULT  15
Q9PL24
ID   Q9PL24      PRELIMINARY;      PRT;   242 AA.
AC   Q9PL24;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL PROTEIN TC0284.
GN   TC0284.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MOPN / NIGG;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Read T.D., Brunham R., Shen C., Gill S.R., Heidelberg J.F., White O.,
RA   Hickey E.K., Peterson J., Umayam L.A., Utterback T., Berry K.,
RA   Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C.,
RA   Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA   Salzberg S.L., Eisen J., Fraser C.M.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE002296; AAF39152.1; -.
DR   TIGR; TC0284; -.
KW   Hypothetical protein.
SQ   SEQUENCE   242 AA;  26914 MW;  0621776E89B3BC84 CRC64;
```

```
  Query Match             8.8%;  Score 76;  DB 2;  Length 242;
  Best Local Similarity   26.2%;  Pred. No. 3;
  Matches   21;  Conservative    6;  Mismatches   21;  Indels   32;  Gaps    3;

Qy      68 PEAYCLCRGCLTGLF---------GEEDVRFRSAPVYMPTVVLRRTPACAGGRSVYT--- 115
           || :| |  |  |             |||: |||  :               |  :||
Db     161 PEPHCNCLHCQIGRIAVEEEDIEVSEEDLTFRSWDI------------SQYGEKMYTVTD 208

Qy     116 --------EAYVTIPVGCTC 127
                   |:  |:||||
Db     209 PLNPEEQFNVYLGTPIGCTC 228
```

```
Search completed: February 16, 2001, 12:52:42
Job time: 197 sec
```

```
Qy   1361 aagtggaacgtgacatctttgccagttgtcagaagaatccaagcaggtattggcttagtt 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    269 AAGTGGAACGTGACATCTTTGCCAGTTGTCAGAAGAATCCAAGCAGGTATTGGCTTAGTT 210

Qy   1421 gtaagggctttaggatcaggctgaatatgaggacaaagtgggccacgttagcatctgcag 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    209 GTAAGGGCTTTAGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATCTGCAG 150

Qy   1481 agatcaatctggaggcttctgtttctgcattctgccacgagagctaggtccttgatcttt 1540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149 AGATCAATCTGGAGGCTTCTGTTTCTGCATTCTGCCACGAGAGCTAGGTCCTTGATCTTT 90

Qy   1541 tctttagattgaaagtctgtctctgaacacaattatttgtaaaagttagtagttctttttt 1600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     89 TCTTTAGATTGAAAGTCTGTCTCTGAACACAATTATTTGTAAAAGTTAGTAGTTCTTTTT 30

Qy   1601 taaatcattaaaagaggcttgctgaa 1626
          |||||||||||||||||||||||||
Db     29 TAAATCATTAAAAGAGGCTTGCTGAA 4
```

```
RESULT   4
HUMCH13C/c
LOCUS       HUMCH13C      296 bp    mRNA            PRI       24-MAR-1994
DEFINITION  Homo sapiens (clone ch13cDNA19B) mRNA sequence.
ACCESSION   L30108
VERSION     L30108.1  GI:463092
KEYWORDS    .
SOURCE      Homo sapiens female 3 month post-natal brain cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 296)
  AUTHORS   Bonaldo,M., Soares,M.-B. and Warburton,D.
  TITLE     Selection of chromosome-specific cDNAs and their corresponding
            genomic clones
  JOURNAL   Unpublished (1993)
FEATURES             Location/Qualifiers
     source          1. .296
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="3 month post-natal"
                     /sex="female"
                     /tissue_type="brain"
                     /map="13q11"
     misc_feature    1. .296
BASE COUNT       84 a     66 c     73 g     73 t
ORIGIN
```

```
  Query Match             17.9%;  Score 294.4;  DB 85;  Length 296;
  Best Local Similarity   99.7%;  Pred. No. 4e-36;
  Matches  295;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1001 acatcgtttttttacttcttctggtagaatttttttaaagcataattggaatccttggataa 1060
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    296 ACATCGTTTTTTTACTTCTTCTGGTAGAATTTGTTAAAGCATAATTGGAATCCTTGGATAA 237

Qy   1061 attttgtagctggtacactctggcctgggtctctgaattcagcctgtcaccgatggctga 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    236 ATTTTGTAGCTGGTACACTCTGGCCTGGGTCTCTGAATTCAGCCTGTCACCGATGGCTGA 177

Qy   1121 ctgatgaaatggacacgtctcatctgacccactcttccttccactgaaggtcttcacggg 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    176 CTGATGAAATGGACACGTCTCATCTGACCCACTCTTCCTTCCACTGAAGGTCTTCACGGG 117

Qy   1181 cctccaggtggaccaaagggatgcacaggcggctcgcatgccccagggccagctaagagt 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    116 CCTCCAGGTGGACCAAAGGGATGCACAGGCGGCTCGCATGCCCCAGGGCCAGCTAAGAGT 57

Qy   1241 tccaaagatctcagatttggttttagtcatgaatacataaacagtctcaaactcgc 1296
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     56 TCCAAAGATCTCAGATTTGGTTTTAGTCATGAATACATAAACAGTCTCAAACTCGC 1
```

(6)

```
RESULT   5
HUMNOTIA/c
LOCUS       HUMNOTIA      323 bp    mRNA            PRI       13-DEC-1993
DEFINITION  Human chromosome specific mRNA.
ACCESSION   L23206
VERSION     L23206.1  GI:434047
KEYWORDS    .
SOURCE      Homo sapiens female 3 month post natal entire brain cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 323)
  AUTHORS   Bonaldo,M., Soares,M.-B. and Warburton,D.
  TITLE     Selection of chromosome-specific cDNAs and their corresponding
            genomic clones
  JOURNAL   Unpublished (1993)
COMMENT     Chromosome 13q11.
FEATURES             Location/Qualifiers
     source          1. .323
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="3 month post natal"
                     /sex="female"
                     /tissue_type="entire brain"
BASE COUNT       41 a     120 c     119 g     43 t
ORIGIN
```

```
  Query Match             15.4%;  Score 252.8;  DB 85;  Length 323;
  Best Local Similarity   99.2%;  Pred. No. 9.9e-30;
  Matches  254;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     21 gccgtgctcagtgccttccaccacacgctgcagctggggccgcgtgagcaggcgcgcaac 80
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    256 GGCGTGCTCAGTGCCTTCCACCACACGCTGCAGCTGGGGCCGCGTGAGCAGGCGCGCAAC 197

Qy     81 gcgagctgcccggcaggggggcaggcccgccgaccgccgcttccggccgcccaccaacctg 140
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    196 GCGAGCTGCCCGGCAGGGGGCAGGCCCGCCGACCGCCGCTTCCGGACGCCCACCAACCTG 137

Qy    141 cgcagcgtgtcgccctgggcctacagaatctcctacgacccggcgaggtaccccaggtac 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 CGCAGCGTGTCGCCCTGGGCCTACAGAATCTCCTACGACCCGGCGAGGTACCCCAGGTAC 77

Qy    201 ctgcctgaagcctactgcctgtgccggggctgcctgaccgggctgttcggcgaggaggac 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 CTGCCTGAAGCCTACTGCCTGTGCCGGGGCTGCCTGACCGGGCTGTTCGGCGAGGAGGAC 17

Qy    261 gtgcgcttccgcagcg 276
          ||||||||||||||||
Db     16 GTGCGCTTCCGCAGCG 1
```

(7)

```
RESULT   6
G31083/c
LOCUS       G31083        250 bp    DNA             STS       31-DEC-1996
DEFINITION  human STS SHGC-19175, sequence tagged site.
ACCESSION   G31083
VERSION     G31083.1  GI:1758484
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 250)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
```